National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

THE UNIVERSITY OF ALBERTA

THE ROLE OF SINGULAR VALUE DECOMPOSITION IN THE ANALYSIS AND DESIGN OF CONTROL SYSTEMS

BY

BRENDA MOUNCE ©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE
IN
PROCESS CONTROL

DEPARTMENT OF CHEMICAL ENGINEERING
EDMONTON, ALBERTA
SPRING 1990

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

ISBN 0-315-60288-0

NL-339 (r. 88/04) c

Canada

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR:	BRENDA MOUNCE
TITLE OF THESIS:	THE ROLE OF SINGULAR VALUE DECOMPOSITION IN THE ANALYSIS AND DESIGN OF CONTROL SYSTEMS
DEGREE:	MASTER OF SCIENCE in PROCESS CONTROL
YEAR DEGREE GRANTED:	1990

Permission is hereby granted to the UNIVERSITY OF ALBERTA to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

Brenda Mounce

#419 - 600 Kirkness Rd.
Edmonton, Alberta
T5Y 2H5

Date: Feb. 12, 1990

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled THE ROLE OF SINGULAR VALUE DECOMPOSITION IN THE ANALYSIS AND DESIGN OF CONTROL SYSTEMS submitted by BRENDA MOUNCE in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE in PROCESS CONTROL.

Dr. S.L. Shah
Supervisor

Dr. D.G. Fisher

R. E. Rink
Dr. R. Rink

Date: ____________

Abstract

Singular value decomposition (SVD) of matrices, both real and complex, has been used as a robust numerical analysis tool in linear algebra for many years. Since approximately 1979, design and analysis tools for process control systems have been developed which utilize SVD. These analysis techniques are well conditioned and numerically robust. However, the outcome of the analysis is dependent on the problem scaling.

There are three main areas where SVD is being used in control system design:

1. multivariable control system design to determine the optimum controlled and manipulated variable pairings for multiloop control schemes, to analyze process interactions, and to determine the number of parameters to retain in model based predictive control techniques such as dynamic matrix control (DMC).

2. model parameter identification and model order reduction techniques which identify a state space model representation of a process from the impulse response. In addition, the order of the model can be reduced through the use of internally balanced state space realizations.

3. robustness and performance analysis of control systems. The condition number and the singular values of the process can give an indication of the robustness and performance characteristics of the system. Recently, the structured singular value was introduced as a measure of robustness in systems with structured uncertainties.

From a survey of the recent literature, recommendations for the use of

these singular value techniques are given. Also, some of the problems with the techniques are highlighted. The use of singular value decomposition in robustness analysis is not covered in this thesis.

Acknowledgements

I would like to express my gratitude to the following individuals:

Dr. S.L. Shah for his guidance, support and suggestions during the course of this study.

Dave Shook for his assistance with MATLAB.

Jim Burkell for his guidance, support and patience while I was completing this thesis.

I would also like to thank the staff of the DACS center for their assistance with the computer equipment and software. Finally, I would like to thank all the others who have assisted me in my work.

TABLE OF CONTENTS

Abstract iv

Acknowledgements vi

List of Tables xi

List of Figures xii

Nomenclature xvii

Chapter 1 Introduction 1

Chapter 2 Singular Value Decomposition 4

2.0 Introduction 4

2.1 Singular Value Decomposition (SVD) 4

2.1.1 Matrix Norms Defined from Singular Values 10

2.2 Applications of SVD in Numerical Linear Algebra 12

2.2.1 Computation of Matrix Rank 12

2.2.2 Condition Number of a Matrix 15

2.2.3 Computation of a Matrix (Pseudo) Inverse 16

2.2.4 Basis Sets for Subspaces of Linear Transformations 17

2.2.5 Angle Between Subspaces 26

2.3 SVD Algorithms 27

2.3.1 Eigenvalue-Eigenvector Routines 27

2.3.2 Golub and Reinsch (1970) SVD Algorithm 28

2.3.3 Adaptive Singular Value Decomposition (ASVD) Algorithm 28

2.4 Summary 30

Chapter 3 **Principal Component Analysis (PCA) in Linear Dynamic Systems** 31

3.0 Introduction 31

3.1 Linear Dynamic System Models 31

3.1.1 Discrete Transfer Function Process Models 32

3.1.1.1 Auto regressive Moving Average (ARMA) Process Models 32

3.1.1.2 Impulse Response (Weighting Sequence) and Step Response Process Models 35

3.1.2 State Space Process Models 40

3.2 PCA of Linear Time Invariant Dynamic Systems 45

3.2.1 Continuous Time Systems 45

3.3 Controllability and Observability Analysis 53

3.3.1 Continuous Time Systems 53

3.3.2 Discrete Time Systems 60

3.4 Summary 64

Chapter 4 **Internally Balanced State Space Process Models** 65

4.0 Introduction 65

4.1 Continuous Time Models 65

4.2 Discrete Time Models 74

4.3 Summary 81

Chapter 5 **Model Identification and Model Order Reduction Techniques Using Singular Value Analysis** 82

5.0 Introduction 82

5.1 Hankel Matrix Model Identification Techniques 84

5.1.1 Ho-Kalman Algorithm and Variations 87

5.2 Model Order Reduction Techniques Using SVD 112

5.2.1 Model Reduction via Truncation of Internally Balanced Realizations 116

5.2.2 Frequency Weighted Truncated Internally Balanced Models 123

5.2.3 Optimal Hankel Norm Model Reduction Techniques 127

5.4 Summary 135

Chapter 6 SVD in the Analysis and Design of Feedback Control Strategies for Multivariable Systems 137

6.0 Introduction 137

6.1 Multiloop Control System Design 140

6.1.1 Process Interactions 140

6.1.2 Singular Value Analysis of Process Interactions 148

6.1.3 Optimum Controlled and Manipulated Variable Pairings for Multiloop Control Schemes Using SVA 152

6.1.4 Process Interaction Measures 159

6.1.4.1 Relative Gain Array (RGA) 160

6.1.4.2 Direct Nyquist Array (DNA) 161

6.1.4.3 IMC Interaction Measure 162

6.1.4.4 Characteristic Loci Measure 163

6.1.4.5 SVA Interaction Measure 167

6.2 SVD in Non-Square System Analysis - Reduction in the Number of Controlled and /or Manipulated Variables 188

6.2.1 Optimum Sensor Location 188

6.2.2 Reduction in the Number of Manipulated Variables 198

6.3 Measures of Controllability and Sensitivity of Multivariable Systems 202

6.4 Summary 210

Chapter 7 Scaling of Process Variables 211

7.0 Introduction 211

7.1 Scaling with Respect to a Reference Value 212

7.2 Empirical Scaling Methods 213

7.3 Normalization and Equilibration 218

7.4 Minimization of Measures of Variability (Semi-Empirical Methods) 220

7.5 Scaling Methods Based on Physical Considerations 221

7.5.1 Dynamic Scaling Procedure (Bonvin, 1987) 221

7.5.2 Physical Scaling Method Based on the Relative Significance of each Variable 226

7.6 Summary 228

Chapter 8 Singular Value Decomposition in Predictive Controller Design 230

8.0 Introduction 230

8.1 Dynamic Matrix Control (DMC) 230

8.2 Principal Components Analysis of DMC 233

8.3 Summary 238

Chapter 9 Conclusions and Recommendations 240

9.1 Conclusions 240

9.1 Recommendations for Further Work 242

References 243

Appendix A Two Interacting Tank System 252

Appendix B Derivation of a Model for a System of Two CSTR's in Series (Bonvin, 1985,1987) 254

Appendix C MATLAB Computer Programs 258

List of Tables

Table

4.1 Condition Number with Respect to Controllability and Observability of the Realizations in Example 3.1 and 4.1 73

6.1 Steady State TFM for the Ethanol-Water Column in Example 6.7 194

6.2 SVD Matrices $(\underline{Z}, \underline{\Sigma}, \underline{V})$ for the Ethanol-Water Column in Example 6 195

6.3 Condition Number and Singular Values for the Steady State Matrix of the Two Interacting Tank System 206

7.1 State Space Realization $(\underline{A},\underline{B},\underline{C},\underline{D})$ and Steady State Gain Matrix for the CSTR System in Example 7.1 214

7.2 SVD Components of the Scaled Steady State Gain Matrix for the CSTR System in Example 7.3 - Geometric Scaling (Row then Column) 219

7.3 SVD Components of the Scaled Steady State Gain Matrix for the CSTR System in Example 7.3 - Geometric Scaling (Column then Row) 219

8.1 Principal Component Magnitudes, $(G_F)_i$, $(G_S)_i$, $\Delta\underline{u}_1$, and Condition Number of the Dynamic Matrix for a Step Change in R for Example 8.1 239

8.2 Principal Component Magnitudes, $(G_F)_i$, $(G_S)_i$, $\Delta\underline{u}_1$, and Condition Number of the Dynamic Matrix for a Step Change in V for Example 8.1 239

List of Figures

Figure

2.1 Singular Vectors for Matrix $\underline{A}$ in Example 2.1 9

2.2 Eigenvectors for the Matrix $\underline{A}$ in Example 2.1 9

2.3 Input Signal Propagation Through the Input and Output Spaces of the 2x2 System in Example 2.5 for an Input $\underline{x} = \underline{e}_1$ 22

2.4 Input Signal Propagation Through the Input and Output Spaces of the 2x2 System in Example 2.5 for an Input $\underline{x} = \underline{v}$ 23

2.5 The Input and Output Subspaces for the Systems Given in Example 2.6. 25

2.6 The Input and Output Subspaces for the Systems Given in Example 2.6. 25

3.1 Impulse Response of a System 37

3.2 Step Response of a System 37

3.3a Unit Impulse Response and Unit Step Response for the Continuous System Given in Equation 3.1.27 44

3.3b Unit Impulse Response and Unit Step Response for the Discrete System Given in Equation 3.1.31 44

3.4 Linear Time Invariant System 45

3.5 Subspaces Representing the Distribution of Energy in the System (the Image of $\underline{G}$) for the System in Example 3.2 51

3.6 State Responses to Test Signals {u(t) and d(t)} (Moore, 1981) 55

5.1a Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters 98

5.1b Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters 98

5.2a Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters 99

5.2b Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters 99

5.3a Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 10 Impulse Response Parameters 101

5.3b Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 10 Impulse Response Parameters 101

5.4a Step Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters 102

5.4b Frequency Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters 102

5.5a Step Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters 103

5.5b Frequency Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters 103

5.6a Step Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 10 Impulse Response Parameters 104

5.6b Frequency Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 10 Impulse Response Parameters 104

5.7a Step Response for 2nd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters 106

5.7b Frequency Response for 2nd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters 106

5.8a Step Response for 2nd Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters 107

5.8b Frequency Response for 2nd Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters 107

5.9a Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters With Noise (Variance ∓0.1) 109

5.9b Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters With Noise (Variance ∓0.1) 109

5.10a Step Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters With Noise (Variance ∓0.1) 110

5.10b Frequency Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters With Noise (Variance ∓0.1) 110

5.11a Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters With Noise (Variance ∓ 0.05) 111

5.11b Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters With Noise (Variance ∓ 0.05) 111

5.12a Step Response for the Reduced Order Models Given in Example 5.3 122

5.12b Frequency Response for the Reduced Order Models Given in Example 5.3 122

5.13a Step Response for the Reduced Order Models Given in Example 5.4 124

5.13b Frequency Response for the Reduced Order Models Given in Example 5.4 124

5.14 Frequency Weighted System 125

6.1 Closed Loop Process Diagram Where ξ = Disturbance Vector (qx1), $\underline{u}$ = Control Input Vector (nx1), $\underline{m}$ = Manipulated Variable Vector, $\underline{e}$ = Feedback Error Vector, $\underline{r}$ = Reference Input Vector, $\underline{y}_{sp}$ = Setpoint Vector (nx1), $\underline{y}$ = Output Vector (nx1) 142

6.2 Closed Loop Signal Transmittances 142

6.3 2x2 Closed Loop System with Multiloop Control Scheme. k_1 and k_2 are the Feedback Controller Gains 147

6.4 A 2x2 Multiloop Control System Showing the SVD of the Process TFM 151

6.5 Optimum Input/Output Pairings Versus Frequency for Example 6.3 (Tung, 1977) 158

6.6a Signals Represented by $\underline{W}_1$ 169

6.6b Signals Represented by $\underline{W}_2$ 169

6.7 Open Loop Step Responses for Two Interacting Tank System (Case a) 176

6.8 Open Loop Step Responses for Two Interacting Tank System (Case b) 176

6.9 Open Loop Step Responses for Two Interacting Tank System (Case a) 177

6.10 Open Loop Step Responses for Two Interacting Tank System (Case a) 177

6.11 Lau's Measure of Interaction for Two Interacting Tank System (Case a) 178

6.12 Lau's Measure of Interaction for Two Interacting Tank System (Case b) 178

6.13 Lau's Measure of Interaction for Two Interacting Tank System (Case c) 179

6.14 Lau's Measure of Interaction for Two Interacting Tank System (Case d) 179

6.15 DNA for Two Interacting Tank System (Case a) 181

6.16 DNA for Two Interacting Tank System (Case d) 181

6.17 DRGA for Two Interacting Tank System (Case d) 182

6.18 Open Loop Step Responses for Example 6.5 (Jensen, 1986) 184

6.19 Lau's Measure of Interaction for Example 6.5 with Loop Pairings u_2-y_1 (σ_1) and u_1-y_2 (σ_2) 184

6.20 DRGA for Example 6.5 (Jensen, 1986) 185

6.21 DNA for Example 6.5 (Jensen, 1986) 185

6.22 Open Loop Step Responses for Example 6.6 (Lau, 1985) 187

6.23 Lau's Measure of Interaction for Example 6.6 (Lau, 1985) 189

6.24 Direct Nyquist Array for Example 6.6 190

6.25 Plot of $|z_1|$ and $|z_2|$ Versus Tray Number for a 50 Tray Ethanol-Water Distillation Column in Example 6.7 (Moore,1986) 196

6.26 Plot of $|z_1|$ - $|z_2|$ Versus Tray Number for a 50 Tray Ethanol-Water Distillation Column in Example 6.7 (Moore, 1986) 196

6.27 Input Signal Propagation Through the Two Interacting Tank System (Appendix I) with R_1=1.0 and R_2=1.0 208

6.28 Input Signal Propagation Through the Two Interacting Tank System (Appendix I) with R_1=0.1 and R_2=1.0 209

7.1 Condition Number and Singular Values of the Steady State Gain Matrix Versus Frequency for the CSTR System Scaled with Respect to Steady State Values (Bonvin, 1985, 1987) 215

7.2 Condition Number and Singular Values of the Steady State Gain Matrix Versus Frequency for the Scaled CSTR System (Bonvin, 1985) 225

Nomenclature

Abbreviations:

ASVD - adaptive singular value decomposition

DMC - dynamic matrix control

GPC - generalized predictive control

Im(A) - image space of linear transformation A

IMC - internal model control

Ker(A) - kernel space of linear transformation A

MIMO - multi-input, multi-output linear system

SISO - single-input, single-output linear system

SVA - singular value analysis

SVD - singular value decomposition

$<\underline{u},\underline{v}>$ - the inner product of the vectors $\underline{u}$ and $\underline{v}$

Variables:

$\underline{\mathrm{A}},\underline{\mathrm{B}},\underline{\mathrm{C}},\underline{\mathrm{D}}$ - continuous state space realization

$\underline{A},\underline{B},\underline{C},\underline{D}$ - discrete state space realization

c_{ij} - cofactors of the return difference matrix $\underline{F}$

dim($\underline{x}$) - dimension of the vector $\underline{x}$

$\underline{e}_i^n$ - i^{th} standard euclidean basis vector where dim($\underline{e}_i$) = n

$\underline{G}_p$ - matrix of transfer functions describing a process

$\underline{G}_p^T$ - the transpose of the real matrix $\underline{G}$

$\underline{G}^{-1}$ - the inverse of the matrix $\underline{G}$

$\underline{G}^{-I}$ - the pseudo-inverse of the matrix $\underline{G}$

$\underline{g}_i(t)$ - the principal component function vector for G(t)

h(i) - i^{th} impulse response parameter

I_n - nxn identity matrix

J - a cost functional

J_F - a cost functional defined with the Frobenius norm

J_s - a cost functional defined with the Spectral norm

M_{ij} - minors of the return difference matrix $\underline{F}$

m - the number of outputs to the process

n - the number of states

p - the number of inputs to the process

P_{ij} - minors of the open loop transfer function matrix

$\mathbb{R}^{m\times n}$ - the set of all real m×n matrices

$\mathbb{R}_r^{m\times n}$ - the set of all real m×n matrices of rank r

R,S,X - real vector spaces

T_s - the sample time

$\underline{v}_i g_i^T$ - the principal component of $\underline{G}$

$\underline{V}$ - unitary matrix (orthogonal) whose columns, $v_i(\underline{G})$, are the right singular vectors of the matrix $\underline{G}$

$\underline{W}^2$ - the Gramian of a general matrix

$\underline{W}_c^2$ - the controllability Gramian

$\underline{W}_o^2$ - the observability Gramian

$\dot{\underline{x}}(t)$ - $d\underline{x}/dt$

$\underline{Z}$ - unitary matrix (orthogonal) whose columns, $z_i(\underline{G})$, are the left singular vectors of the matrix $\underline{G}$

$\underline{0}$ - the zero matrix

ϵ - machine precision

$\gamma(\underline{G})$ - condition number of the matrix $\underline{G}$

ω - angular frequency

$\underline{\Sigma}$ - a diagonal matrix whose diagonal elements are the singular values of the matrix $\underline{G}$, $\sigma_i(\underline{G})$

$\sigma_i(\underline{G})$ - i^{th} singular value of the matrix $\underline{G}$ (or component magnitude)

$\sigma_{min}(\underline{G})$ - minimum singular value of the matrix $\underline{G}$

$\sigma_{max}(\underline{G})$ - maximum singular value of the matrix $\underline{G}$

$\underline{\Gamma}$ - infinite impulse hankel matrix

$\underline{\Gamma}_{N,N}$ - NxN impulse hankel matrix

Γ - hankel operator for continuous systems

Vector and Matrix Norms:

Euclidean vector norm:

$$\left\| \underline{x} \right\| = (\underline{x}^T \underline{x})^{1/2} = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij}^2 \right)^{1/2}$$

Spectral matrix norm:

$$\left\| \underline{G} \right\|^2 = \max_{\|x\| \neq 0} \frac{\left\| \underline{G}x \right\|^2}{\left\| x \right\|^2} = \max_i \lambda_i \left(\underline{G}^T \cdot \underline{G} \right) = \sigma_{max}^2(\underline{G})$$

Frobenius matrix norm:

$$\left\| \underline{G} \right\|_F = \left(\sum_{i=1}^{n} \sigma_i^2(\underline{G}) \right)^{1/2} = \text{trace } \left(\underline{G} \cdot \underline{G}^T \right)^{1/2}$$

Hankel matrix norm:

$$\left\| \underline{G} \right\|_H = \sigma_{max}\left(\underline{\Gamma}\left(\underline{G}_p \right) \right) = \lambda_{max}^{1/2} \left(\underline{W}_c^2 \cdot \underline{W}_o^2 \right)$$

Induced matrix 1-norm:

$$\left\| \underline{G} \right\|_{i1} = \max_i \sum_{j=1}^{m} |g_{ij}| \quad \text{(max. column sum)}$$

Induced matrix ∞-norm: $$\left\|\underline{G}\right\|_{i\infty} = \max_{j} \sum_{i=1}^{n} |g_{ij}| \quad \text{(max. row sum)}$$

L^2 matrix norm: $$\left\|\underline{G}\right\|_2^2 = \int_{i=1}^{\infty} \text{trace}\left(\underline{A}^T \underline{A}\right)\, dt$$

$$= \sigma_{\max}\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \underline{G}^T(e^{i\omega}) \cdot \underline{G}(e^{i\omega})\, d\omega \right)$$

Chapter 1 INTRODUCTION

The objective of this thesis was to survey the literature and prepare a tutorial review and evaluation of various techniques which employ singular value decomposition in the analysis and design of process control systems.

Singular value decomposition, SVD, of real and complex matrices has been used as a numerical analysis tool for several years (Browne, 1930, Eckart, 1939, Forsythe, 1977, Golub, 1970, Klema 1980). SVD provides a reliable and robust technique for calculating the rank, condition number, and the inverse (or pseudo-inverse) of a matrix. A diagonal matrix containing the singular values and two unitary matrices whose columns are the singular vectors are obtained from the SVD of a matrix. There are very reliable and efficient algorithms available which will decompose matrices into their singular value and singular vector matrices.

In recent years, several researchers have been investigating the use of SVD as a computational tool for analysis and design of process control systems (Moore, 1979a, 1981, Glover, 1984, Doyle, 1979, Maurath, 1985, Kung, 1978). There are four main categories of process control system design and analysis research where SVD derived analysis techniques can be employed: multi-input/multi-output control system design, process parameter identification, process model order reduction and stability and robustness analysis. The techniques based on SVD of system matrices are always numerically well conditioned due to the numerically well conditioned nature of the decomposition matrices. However, the singular value analysis techniques are affected by the scaling of the system variables and/or matrices. Several researchers have proposed methods of system scaling to ensure that the physical significance of the problem has not been changed by the scaling procedure (Lau, 1985, Keller, 1987, Bonvin, 1985, 1987, Bryant, 1983, Johnston, 1984, 1985). If the systems are scaled in a reasonable manner based on the problem being investigated, the singular value analysis techniques should provide "good" consistent results.

In this thesis, the application of SVD techniques for process model identification and order reduction and for multiloop and multivariable control system design was reviewed and evaluated. It has been shown how singular value decomposition of a system transfer function matrix can be used to gain a better understanding of process interactions. Throughout this work, an effort has been made to clarify the application and understanding of these techniques.

This thesis is organized as follows. Chapter 2 of this thesis introduces the principles and properties of singular value decomposition. A discussion of principal component analysis of linear dynamic systems, which forms the basis for the singular value analysis techniques discussed in the remainder of this thesis, is presented in Chapter 3. As well, the use of SVD to analyze the controllability and observability of linear dynamic systems is discussed in Chapter 3. In Chapter 4, the concept of internally balanced state space realizations of process models is introduced. These balanced models are used in several of the model identification and reduction techniques discussed later in the thesis. Chapter 5 details SVD process model identification and process model order reduction techniques which employ SVD. Chapter 6 examines the use of singular value analysis (SVA) techniques in the design of MIMO control strategies covering "optimal" input/output pairings for multiloop control strategies and interaction analysis. The primary disadvantage of SVD techniques in control system analysis, the scaling of the input and output variables, is discussed in

Chapter 7. Chapter 8 investigates the application of SVD in predictive controller design. The conclusions and recommendations of this thesis are presented in Chapter 9.

Throughout this thesis, examples are used to illustrate the principles of SVD as they apply to system analysis and design. All the examples were generated using the software package PC-MATLAB (The MathWorks Inc.) whose basic data element is a matrix. PC-MATLAB is an interactive software system written in the C language which provides graphic representations, programmable macros, IEEE arithmetic, a fast interpreter and many analytical commands (Moler, 1986). An IBM compatible AT personal computer with a math co-processor was used for all the computations. SVD is an internal program within MATLAB which employs the algorithm of Golub and Reinsch (1970).

Although the use of SVD in the analysis of robust performance and robust stability of multivariable systems has been investigated by several researchers (Doyle, 1979, 1981, Johnston, 1987, Arkun, 1984, Skogestad, 1987, Morari, 1983, 1985, 1989), this topic was not included as part of this thesis.

Chapter 2 Singular Value Decomposition

2.0 Introduction

Although mathematicians have been using singular value decomposition (SVD) as a numerical tool for several years, it is only recently that its application in the field of process control has been explored. This chapter discusses some of the numerical properties of singular value decomposition of matrices as well as applications of SVD in linear algebra. One of the useful properties which has been exploited for control system analysis is the use of SVD to analyze linear transformations and solve the corresponding linear equations.

2.1 Singular Value Decomposition (SVD)

Singular value decomposition of a matrix is an efficient and reliable method of determining the rank, condition number and (pseudo) inverse of matrices. Given a real matrix $\underline{\mathbf{A}} \in \mathbf{R}^{nxm}$ where n and m are arbitrary integer numbers, there will exist orthonormal matrices $\underline{\mathbf{Z}} \in \mathbf{R}^{nxn}$ and $\underline{\mathbf{V}} \in \mathbf{R}^{mxm}$ such that

$$\underline{\mathbf{A}} = \underline{\mathbf{Z}}\,\underline{\mathbf{\Sigma}}\,\underline{\mathbf{V}}^{T} \qquad [2.1.1]$$

where

$$\underline{\mathbf{\Sigma}} = \begin{bmatrix} \underline{\mathbf{S}} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \underline{\mathbf{0}} \end{bmatrix} \qquad [2.1.2]$$

$\underline{\mathbf{\Sigma}} \in \mathbf{R}^{nxm}$ and $\underline{\mathbf{S}} = \mathrm{diag}\ \{\sigma_1, \sigma_2, \ldots, \sigma_r\}$ (Eckart, 1939, Klema, 1980). A SVD theorem also exists for complex matrices. If $\underline{\mathbf{A}}$ is a full rank matrix and m = n,

$$\underline{\Sigma} = \text{diag}\ \{\sigma_1,\sigma_2,\cdots,\sigma_r,\ \sigma_{r+1},\cdots,\sigma_n\} \qquad [2.1.3]$$

An orthonormal matrix satisfies the following property:

$$\underline{A}^T\underline{A} = \underline{A}\underline{A}^T = \underline{I} \qquad [2.1.4]$$

The columns of the matrices $\underline{Z}$ and $\underline{V}$ are defined as the left and right singular vectors, respectively, of the matrix $\underline{A}$. The left and right singular vectors are the orthonormal eigenvectors of $\underline{A}\underline{A}^T$ and $\underline{A}^T\underline{A}$, respectively. The diagonal elements of the matrix $\underline{\Sigma}$ are defined as the singular values of the matrix $\underline{A}$ and are the non-negative square roots of the eigenvalues of $\underline{A}^T\underline{A}$ or $\underline{A}\underline{A}^T$. The singular values are arranged such that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq \sigma_{r+1} \geq \cdots \geq \sigma_n \geq 0 \qquad [2.1.5]$$

Since $\underline{A}\underline{A}^T$ and $\underline{A}^T\underline{A}$ are symmetric matrices, all the singular values of $\underline{A}$ will be real and non-negative. Symmetric matrices have the following property:

$$\underline{A}^T = \underline{A} \qquad [2.1.6]$$

In dyadic form, the singular value decomposition of a (nxm) matrix can be written as:

$$\underline{A} = \sum_{i=1}^{n} \sigma_i \underline{z}_i \underline{v}_i^T \qquad [2.1.7]$$

where

$$\underline{z}_i = i^{th} \text{ column of } \underline{Z}$$

$$\underline{v}_i = i^{th} \text{ column of } \underline{V}$$

The SVD thereom for the special case of square matrices with real elements has been proven by Sylvester (1889). Corresponding results for complex, square matrices were proved by L. Autonne (1902) and by Browne (1930). In 1939, Eckart and Young proved the theroem for general rectangular matrices.

Singular values and singular vectors of a matrix can be viewed as generalized eigenvalues and eigenvectors and can be utilized in a similar manner. The linear equation

$$\underline{y} = \underline{A}\cdot\underline{x} \qquad \text{[2.1.8]}$$

represents a linear transformation in which the vector $\underline{x} \in \mathbb{R}^n$ is transformed into a new vector $\underline{y} \in \mathbb{R}^n$ via $\underline{A} \in \mathbb{R}^{nxn}$. A set of vectors, $\underline{y}$, exists for which the vector $\underline{x}$ is transformed into a multiple of itself such that:

$$\underline{Ax} = \lambda\underline{x} = \underline{y} \qquad \text{[2.1.9a]}$$

$$(\underline{A} - \lambda\underline{I})\underline{x} = \underline{0} \qquad \text{[2.1.9b]}$$

The eigenvalues of $\underline{A}$ are defined as the set of non-zero values, $\{\lambda_i\text{:}i=1,2,...,n\}$, such that the set of n homogeneous equations in m unknowns, given in equations 2.1.9b has a non-trivial solution (Wilkinson, 1965). Therefore, the set of eigenvalues are the roots of the characteristic equation

$$\det|\underline{A} - \lambda\underline{I}| = \underline{0} \qquad \text{[2.1.10]}$$

since $\underline{x}\neq\underline{0}$. The vector solutions $(\underline{x}_i\text{: } i=1,2,..,n)$ to the n homogeneous equations,

$$\underline{Ax}_i = \lambda_i\underline{x}_i \qquad \text{[2.1.11]}$$

are defined as the right hand eigenvectors of $\underline{A}$. The set of left hand eigenvectors of $\underline{A}$ $(\underline{y}_i\text{: } i=1,2,..,n)$ are the set of solutions to the linear system

$$\underline{y}_i^T\underline{A} = \lambda_i\underline{y}_i^T \qquad \text{[2.1.12]}$$

whereby $\underline{y}$ is transformed into a multiple of itself. The set of eigenvectors is generally normalized giving a set of vectors of unit norm (MacFarlane, 1970).

The eigenvalues of $\underline{A}$ are related to the singular values since

$$\sigma_{\min}(\underline{A}) \leq |\lambda_i(\underline{A})| \leq \sigma_{\max}(\underline{A}) \qquad 1\leq i\leq n \qquad \text{[2.1.13]}$$

If the matrix $\underline{A} \in \mathbb{R}^{nxn}$ is symmetric with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$ then,

$$\sigma_i(A) = |\lambda_i(A)| \qquad [2.1.14]$$

and the right and left singular vectors are the same and equivalent to the eigenvectors (Wilkinson, 1965).

Eigenvalues and eigenvectors can be used to transform a given square matrix into a diagonal matrix which can be useful in the solution of systems of linear equations (Boyce and DiPrima, 1977 pp. 292-297). A matrix, $\underline{T}$, can be formed where each column is an eigenvector of $\underline{A}$. If $\underline{A} \in \mathbb{R}^{nxn}$ has n distinct eigenvalues, there will be n unique and linearly independent eigenvectors. The inverse of $\underline{T}$, $\underline{T}^{-1}$, will exist and

$$\underline{T}^{-1}\underline{A}\underline{T} = \underline{D} \qquad [2.1.15]$$

where $\underline{D} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ (Boyce and DiPrima, 1977). If there are multiple eigenvalues of $\underline{A}$, the set of eigenvectors will contain linearly dependant vectors and the inverse of $\underline{T}$ will not exist. Therefore, a unique solution to equation 2.1.15 will not exist.

Similarly, singular vectors and singular values of the matrix $\underline{A}$ can be used to transform $\underline{A}$ into a diagonal matrix. If $\underline{A} \in \mathbb{R}^{nxn}$ is a symmetric matrix, the left and right singular vectors corresponding to a singular value will be the same such that $\underline{Z} = \underline{V}$. From equation 2.1.1,

$$\underline{Z}^T\underline{A}\underline{V} = \underline{Z}^T\underline{A}\underline{Z} = \underline{V}\underline{A}\underline{V}^T = \underline{\Sigma} \qquad [2.1.16]$$

where $\underline{\Sigma}$ is a diagonal matrix with the singular values on the diagonal. If $\underline{A}$ is not symmetric, premultiplication by the left singular vector transposed and postmultiplication by the right singular vector will always generate a diagonal matrix $\underline{\Sigma}$. The inverse of $\underline{Z}$ and $\underline{V}$, which is equivalent to the transpose, will always exist even if $\underline{A}$ has repeated singular values since $\underline{Z}$ and $\underline{V}$ are always orthonormal by definition. Since the singular vectors are orthonormal by definition, they will span the subspace for the linear transformation defined by the matrix $\underline{A}$ and will form a basis set for that subspace. On the other hand, eigenvectors of $\underline{A}$ may not be orthogonal.

Example 2.1 (Boyce, 1977): If

$$\underline{A} = \begin{bmatrix} 5 & -1 \\ 3 & 1 \end{bmatrix} \qquad [2.1.17]$$

the singular value decomposition of $\underline{A}$ gives,

$$\underline{Z} = \begin{bmatrix} 0.865 & 0.502 \\ 0.502 & -0.865 \end{bmatrix} \quad \underline{\Sigma} = \begin{bmatrix} 5.84 & 0 \\ 0 & 1.37 \end{bmatrix} \quad \underline{V} = \begin{bmatrix} 0.998 & -0.062 \\ -0.062 & -0.998 \end{bmatrix} \qquad [2.1.18]$$

The eigenvectors and the corresponding eigenvalues of $\underline{A}$ are

$$\underline{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \lambda_1 = 4.0$$
$$\underline{x}_2 = \begin{bmatrix} 0.33 \\ 1 \end{bmatrix} \qquad \lambda_2 = 2.0 \qquad [2.1.19]$$

As illustrated in Figure 2.1, the set of singular vectors form an orthonormal basis set for the subspace of the linear transformation A. However, Figure 2.2 shows that even though the set of associated eigenvectors are linearily independant and span the subspace, they are not orthogonal. It should be noted that $\sigma_1 < |\lambda_i| < \sigma_2$.

The matrices, ($\underline{Z}$,$\underline{V}$,$\underline{\Sigma}$), from the singular value decomposition of $\underline{A}$ are numerically very well conditioned. $\underline{Z}$ and $\underline{V}$ are orthonormal matrices which always have a condition number of 1.0 and $\underline{\Sigma}$ is diagonal. If the matrix $\underline{A}$ is real, then the diagonal elements of $\underline{\Sigma}$ will be real and non-negative. $\underline{Z}$ and $\underline{V}$ will be real for all matrices: square and non-square, full rank and rank deficient. In contrast, eigenvectors and eigenvalues can be real and/or complex. There is no guarantee that the resulting eigenvectors will be

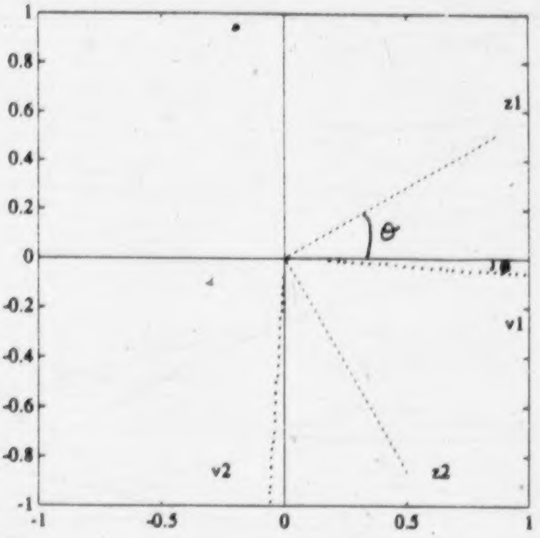

Figure 2.1: Singular Vectors of the Matrix $\underline{A}$ in Example 2.1

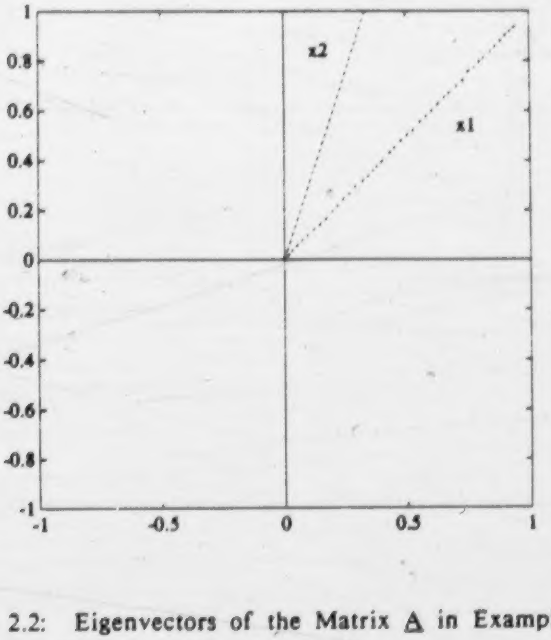

Figure 2.2: Eigenvectors of the Matrix $\underline{A}$ in Example 2.1

orthogonal or independent. Therefore, the use of SVD as a computational tool will always provide an environment with "good" numerical properties. The usefulness of eigenvectors and eigenvalues as analysis tools are dependant on the matrix describing the linear transformation.

Example 2.2 (Morari, 1989): If

$$\underline{A} = \begin{bmatrix} 0.871 & -1.320 \\ 1.578 & -0.095 \end{bmatrix} \qquad [2.1.20]$$

the singular value decomposition of $\underline{A}$ gives,

$$\underline{Z} = \begin{bmatrix} 0.707 & 0.707 \\ 0.707 & -0.707 \end{bmatrix} \quad \underline{\Sigma} = \begin{bmatrix} 2.00 & 0 \\ 0 & 1.00 \end{bmatrix} \quad \underline{V} = \begin{bmatrix} 0.866 & -0.50 \\ -0.50 & -0.866 \end{bmatrix} \qquad [2.1.21]$$

The eigenvectors and the corresponding eigenvalues of $\underline{A}$ are

$$\underline{x}_1 = \begin{bmatrix} 0.306 + 0.862i \\ 1.00 \end{bmatrix} \qquad \lambda_1 = 0.388 + 1.369i$$
$$\underline{x}_2 = \begin{bmatrix} 0.306 - 0.862i \\ 1.00 \end{bmatrix} \qquad \lambda_2 = 0.388 - 1.360i \qquad [2.1.22]$$

Although the matrix is real, its eigenvalues and eigenvectors are complex. On the other hand, the singular values and vectors are real. If the matrix represents a linear transformation of a real system, it will be easier to analyze the transformation in the subspace defined by the singular vectors rather than the subspace defined by the eigenvectors.

2.1.1. Matrix Norms Defined From Singular Values

Matrix norms are used to indicate the "size" of a matrix in the same manner as a vector norm measures the length of a vector. The maximum singular value of a matrix defines a unitarily invariant matrix norm, the spectral norm, which is similar in definition to the Euclidean norm of a vector (Klema, 1980). A unitarily invariant norm is one in which multiplication of a matrix by a unitary matrix will not change the value of the norm of the original matrix.

The spectral norm of a matrix $\underline{A} \in \mathbb{R}^{nxm}$ is defined as (Klema, 1980):

$$A_2 = \max_{\underline{x} \neq 0} \frac{\underline{A}\cdot\underline{x}_2}{\underline{x}_2} = \sigma_{max}(\underline{A}) \qquad [2.1.23]$$

If $\underline{A}$ is invertible,

$$\underline{A}^{-1}{}_2 = 1/\sigma_{min}(\underline{A}) = \sigma_{max}(\underline{A}^{-1}) \qquad [2.1.24]$$

The spectral norm has several properties which are characteristic of all matrix norms (Maurath, 1987, Klema, 1980, Forsythe, 1977):

i) if $\underline{A} \neq \underline{0}$, then $\underline{A}_2 > 0$ [2.1.25a]

ii) if a is a scalar, then $a\underline{A}_2 = a \cdot \underline{A}_2$ [2.1.25b]

iii) $\underline{A}_1 + \underline{A}_2{}_2 \leq \underline{A}_1{}_2 + \underline{A}_2{}_2$ [2.1.25c]

iv) $\underline{A}_2 = \underline{\Sigma}_2$ [2.1.25d]

v) $\underline{Ax}_2 \leq \underline{A}_2 \cdot \underline{x}_2 = \sigma_1(\underline{A}) \cdot \underline{x}_2$ [2.1.25e]

vi) $\sigma_{min}(\underline{A}) \leq \dfrac{\underline{A}x_2}{\underline{x}_2} \leq \sigma_{max}(\underline{A})$ [2.1.25f]

Another unitarily invariant norm, the Frobenius norm, can be defined from the singular values of a matrix:

$$\underline{A}_F^2 = \text{trace}\,(\underline{A}^T\underline{A}) = \sum_{i=1}^{n} \sigma_i^2(\underline{A}) \qquad [2.1.26]$$

There exists a relationship between the spectral norm and the Frobenius norm (Morari, 1989):

$$\underline{A}_2 \leq \underline{A}_F \leq \sqrt{n}\ \underline{A}_2 \qquad [2.1.27]$$

where n is the number of nonzero singular values of $\underline{A}$.

2.2 Applications of SVD in Numerical Linear Algebra

2.2.1 Computation of Matrix Rank

It has been generally accepted in the literature that SVD is the most reliable method of determining the numerical rank of a general matrix (Klema, 1980). If $\underline{A} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T$, the rank of $\underline{A}$ will be the same as the rank of $\underline{\Sigma}$ since $\underline{Z}$ and $\underline{V}$ are always of full rank. The number of non-zero singular values of $\underline{A}$ such that $\sigma_i \geq \epsilon$ where ϵ is the machine precision determines the rank of $\underline{\Sigma}$. All the singular values of $\underline{A}$ will be non-zero if and only if the columns of $\underline{A}$ are linearly independent. The minimum singular value of a matrix, $\sigma_n > 0$, indicates how close a matrix is to singularity and matrices of lesser rank (Klema, 1980). Therefore, singular values can be used to determine linear dependence of rows and columns in a matrix corresponding to the equations and the variables, respectively, in a set of linear equations.

In a practical application, there is generally noise or data errors associated with the elements in the matrix of interest. Using conventional matrix elimination methods to calculate the rank, these data errors can result in a matrix which appears to have full numerical rank (n) although the actual rank is some number r<n. Using SVD, the actual rank can be determined if an estimate of the data errors is known. Given a matrix $\underline{A} \in \mathbb{R}^{nxn}$ where

$$\underline{A} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T \qquad [2.2.1]$$

with singular values given in equation 2.1.3, a matrix $\underline{B}$ with rank r can be defined as

$$\underline{B} = \underline{Z}\cdot\hat{\underline{\Sigma}}\cdot\underline{V}^T \qquad [2.2.2]$$

where $\hat{\underline{\Sigma}} = \mathrm{diag}(\sigma_1,\cdots,\sigma_r,0,0,\cdots,0)$. If $\underline{B}$ is used to approximate $\underline{A}$, the approximation error in terms of a spectral norm is given by:

$$\begin{aligned}\left\|\underline{A}-\underline{B}\right\|_2 &= \left\|\underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T - \underline{Z}\cdot\underline{\hat{\Sigma}}\cdot\underline{V}^T\right\|_2 \\ &= \left\|\underline{Z}\cdot(\underline{\Sigma} - \underline{\hat{\Sigma}})\cdot\underline{V}^T\right\|_2 \\ &= \left\|\underline{\Sigma} - \underline{\hat{\Sigma}}\right\|_2 = \sigma_{r+1}\end{aligned} \qquad [2.2.3]$$

since $\underline{Z}$ and $\underline{V}$ are unitary matrices. Similarly,

$$\left\|\underline{A}\right\|_F - \left\|\underline{B}\right\|_F \leq \left[\sum_{i=r+1}^{n} \sigma_i^2\right]^{1/2} \qquad [2.2.4]$$

where $\|\cdot\|_F$ is the Frobenius norm (Klema, 1980). If a "zero threshold" or noise criterion, δ, is chosen such that $\sigma_{r+1}<\delta$, the matrix $\underline{A}$ will be equivalent to the matrix $\underline{B}$ within the noise band (Klema, 1980). The magnitude of δ will be determined by the uncertainty in the data. In this case, $\underline{A}$ is said to have numerical rank r.

When a perturbation is added to a matrix, it can be shown that the singular values of the matrix will change by no more than the size of the perturbation. Given a matrix $\underline{A} \in \mathbb{R}^{nxm}$ with singular values $\{\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \geq \cdots \geq \sigma_n \geq 0\}$, if $\underline{x}$ varies throughout the vector space, $\mathbb{R}^k$, which is the subspace spanned by the singular vectors $\{\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_k\}$, the singular values can be expressed as the local minimum of the spectral norm such that (Wilkinson, 1965, MacFarlane, 1979):

$$\sigma_k(\underline{A}) = \min_{d(\delta)=n-k+1} \; \max_{\substack{\underline{x}\in\delta \\ \underline{x}\neq 0}} \frac{\left\|\underline{A}\cdot\underline{x}\right\|_2}{\left\|\underline{x}\right\|_2} \qquad [2.2.5]$$

If $\underline{E}$ is a perturbation matrix of $\underline{A}$ such that $\underline{A}+\underline{E}$ has singular values $\{\tau_1 \geq \tau_2 \geq \ldots \geq \tau_n \geq 0\}$, then

$$\left|\sigma_i - \tau_i\right| \leq \left\|\underline{E}\right\|_2 \qquad 1 \leq i \leq n \qquad [2.2.6]$$

Example 2.3: If the noise criterion is chosen as $\delta \leq 0.0001$ for the linear system with matrix $\underline{A}$ given by,

$$\underline{A} = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & -1 \\ 2 & 1.999 & -3 \end{bmatrix} \qquad [2.2.7a]$$

with singular values

$$\underline{\Sigma}_{\underline{A}} = \begin{bmatrix} 5.152 & 0 & 0 \\ 0 & 2.731 & 0 \\ 0 & 0 & 0.0001 \end{bmatrix} \qquad [2.2.7b]$$

a matrix $\underline{B}$, obtained from equation 2.2.2, will be given by

$$\underline{B} = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & -1 \\ 2 & 2 & -3 \end{bmatrix} \qquad [2.2.8a]$$

with singular values

$$\underline{\Sigma}_{\underline{B}} = \begin{bmatrix} 5.152 & 0 & 0 \\ 0 & 2.731 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad [2.2.8b]$$

For this system,

$$\left\| \underline{A} - \underline{B} \right\|_2 = 0.0001 = \sigma_3 \qquad [2.2.9]$$

Therefore, the actual rank of $\underline{A}$ is 2.

Example 2.4: If the elements in the matrix $\underline{A}$, given in equation 2.1.17, are perturbed by 1%, the resulting perturbed matrix will be,

$$\begin{aligned} \underline{E} &= \begin{bmatrix} 0.009 & -0.013 \\ 0.016 & -0.001 \end{bmatrix} \\ \underline{A} + \underline{E} &= \begin{bmatrix} 0.880 & -1.333 \\ 1.594 & -0.096 \end{bmatrix} \end{aligned} \qquad [2.2.10]$$

where

$$\underline{\Sigma}_{\underline{E}} = \begin{bmatrix} 0.020 & 0 \\ 0 & 0.010 \end{bmatrix} \qquad [2.2.11a]$$

$$\underline{\Sigma}_{\underline{A}+\underline{E}} = \begin{bmatrix} 2.020 & 0 \\ 0 & 1.010 \end{bmatrix} \qquad [2.2.11b]$$

are the singular value matrices of $\underline{E}$ and $\underline{A}+\underline{E}$, respectively. Therefore,

$$|\sigma_1 - \tau_1| = 0.02 = \|\underline{E}\|_2 \qquad [2.2.12a]$$

and

$$|\sigma_2 - \tau_2| = 0.01 < \|\underline{E}\|_2 \qquad [2.2.12b]$$

2.2.2 Condition Number of a Matrix

The condition number of a matrix, $\gamma(\underline{A})$, is defined as:

$$\gamma(\underline{A}) = \|\underline{A}\|_2 \cdot \|\underline{A}^{-1}\|_2 = \sigma_{max}(\underline{A})\cdot\sigma^{-1}_{min}(\underline{A}) \geq 1 \qquad [2.2.13]$$

where $\underline{A} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T$. The condition number indicates potential linear dependencies of columns in the matrix and the approach to singularity of the matrix. If $\gamma(\underline{A}) = 1$, $\underline{A}$ is an orthonormal matrix whose columns are linearly independent. If $\gamma(\underline{A})$ is infinite, $\underline{A}$ is rank deficient. In this case, one or more of the columns of $\underline{A}$ will be linearly dependant. If $\gamma(\underline{A})$ is large, there are some columns in $\underline{A}$ which are nearly linearly dependant and $\underline{A}$ will be nearly rank deficient (Klema, 1980).

The condition number of $\underline{A}$ can be used to analyze the solution set for the linear equation given in equation 2.1.8,

$$\underline{y} = \underline{A}\cdot\underline{x}$$

The accuracy of the computed solution set, $\underline{x}$, is indicated by the condition number of $\underline{A}$ since the condition number is a measure of the magnification of errors in $\underline{A}$ and/or $\underline{y}$ in the solution $\underline{x}$. In the case where $\underline{A}$ has a large $\gamma(\underline{A})$, the solution to the linear problem will be very sensitive to changes in the matrix $\underline{A}$ and errors in the data sets. The calculated solution $\underline{x}_\Delta$ can be far from the true solution $\underline{x}$. If the perturbed system is given by,

$$\underline{y} + \Delta\underline{y} = \underline{y}_\Delta = (\underline{A} + \Delta\underline{A})\cdot\underline{x}_\Delta \qquad [2.2.14]$$

where $\underline{x}_\Delta$ is the solution vector for the perturbed system, Moore (1979a) has shown that

$$\frac{\|\Delta\underline{x}\|_2}{\|\underline{x}\|_2} = \gamma(\underline{A}_\Delta)\cdot\frac{\|\Delta\underline{y}\|_2}{\|\underline{y}\|_2} \tag*{[2.2.15]}$$

2.2.3 Computation of a Matrix (Pseudo) Inverse

SVD is a reliable and easy method of computing a matrix inverse if the matrix has full rank or a pseudo-inverse for rank deficient matrices. The matrix $\underline{X}$ will be a pseudo-inverse of the matrix $\underline{A}$ if (Golub, 1970, Barnett, 1971):

$$\text{a)}\ \ \underline{AXA} = \underline{A} \tag*{[2.2.16a]}$$

$$\text{b)}\ \ \underline{XAX} = \underline{X} \tag*{[2.2.16b]}$$

$$\text{c)}\ \ (\underline{AX})^T = \underline{AX} \tag*{[2.2.16c]}$$

$$\text{d)}\ \ (\underline{XA})^T = \underline{XA} \tag*{[2.2.16d]}$$

The Moore-Penrose generalized inverse or pseudo-inverse for $\underline{A} \in \mathbb{R}^{nxm}$ $(m \leq n)$ is defined as: (Barnett, 1971)

$$\underline{A}^{-I} = (\underline{A}^T\underline{A})^{-1}\underline{A}^T \tag*{[2.2.17]}$$

If $\underline{A}$ has full rank, $\underline{A}$ will be invertible and $\underline{A}^{-I} = \underline{A}^{-1}$. If

$$\underline{A} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T \tag*{[2.2.18]}$$

the Moore-Penrose generalized inverse is given by

$$\underline{A}^{-I} = \underline{V}\cdot\underline{\Sigma}^{-1}\cdot\underline{Z}^T \tag*{[2.2.19a]}$$

where

$$\underline{\Sigma}^{-1} = \text{diag}\ (\sigma_i^{-1}) \tag*{[2.2.19b]}$$

with

$$\sigma_i^{-1} = \begin{cases} 1/\sigma_i & \text{if } \sigma_i > 0 \quad \text{for } 1 \leq i \leq n \\ 0 & \text{if } \sigma_i = 0 \quad \text{for } 1 \leq i \leq n \end{cases} \tag*{[2.2.19c]}$$

The pseudo-inverse is useful when solving linear least squares problems with rank deficient matrices to obtain a minimal norm or least squares solution. Given the linear problem in equation 2.1.8 where $\underline{y} \in \mathbb{R}^n$ and $\underline{A} \in \mathbb{R}^{n\mathrm{x}m}$, a unique solution,

$$\underline{x} = \underline{A}^{-1}\underline{y} \in \mathbb{R}^m \qquad [2.2.20]$$

can be found only if the matrix $\underline{A}$ has full rank and $m \leq n$. In this case, the norm of the solution error, $\underline{y}-\underline{A}\underline{x}\ _2$, will be zero. If the matrix $\underline{A}$ has rank r which is less than n, there will be an infinite number of solutions $\underline{x}$ to equation 2.1.8. An approximate solution

$$\hat{\underline{x}} = \underline{A}^{-I}\underline{y} = (\underline{A}^T\underline{A})^{-1}\underline{A}^T\underline{y} \in \mathbb{R}^m \qquad [2.2.21]$$

can be found such that the solution error $\underline{y}-\underline{A}\hat{\underline{x}}\ _2$ is minimal in the least squares sense (Golub, 1970, Maurath, 1985).

2.2.4 Basis Sets for Subspaces of Linear Transformations

Klema (1980) presents a review of the numerical applications of SVD with respect to linear transformations. SVD is a very reliable numerical tool for calculating orthonormal basis vector sets for the four fundamental subspaces of a linear transformation. For the linear system described by equation 2.1.8,

$$\underline{y} = \underline{A}\cdot\underline{x}$$

$\underline{A} \in \mathbb{R}^{n\mathrm{x}m}$ is a matrix which represents the linear transformation A: $X \Rightarrow Y$. X and Y are real vector spaces with $\dim(X) = m$ and $\dim(Y) = n$ such that $\underline{y} \in \mathbb{R}^n$ and $\underline{x} \in \mathbb{R}^m$. There are four fundamental subspaces which characterize the linear transformation (Moore, 1979a):

$$1.\ \mathrm{Im}(\underline{A}) \qquad [2.2.22a]$$

$$2.\ \mathrm{Im}(\underline{A}^{-1}) = \mathrm{Ker}(\underline{A})^{\perp} \qquad [2.2.22b]$$

$$3.\ \mathrm{Ker}(\underline{A}) \qquad [2.2.22c]$$

$$4.\ \mathrm{Ker}(\underline{A}^{-1}) = \mathrm{Im}(\underline{A})^{\perp} \qquad [2.2.22d]$$

$\mathrm{Im}(\underline{A})$ or image of the transformation A is the subspace comprised of the set of all possible vectors $\underline{y}$ such that equation 2.1.8 has a unique solution. $\mathrm{Ker}(\underline{A})$ or kernel of the transformation A is the subspace comprised of the set of all possible solution vectors $\underline{x}$ to the homogeneous equation

$$\underline{0} = \underline{A}\cdot\underline{x} \qquad [2.2.23]$$

$\mathrm{Im}(\underline{A})^{\perp}$ and $\mathrm{Ker}(\underline{A})^{\perp}$ are the projections of the subspace of the linear transformation A orthogonal to the subspaces $\mathrm{Im}(\underline{A})$ and $\mathrm{Ker}(\underline{A})$, respectively.

The linear equation 2.1.8 can be written as

$$\underline{E}_n\cdot\underline{y} = \underline{A}\cdot\underline{E}_m\cdot\underline{x} \qquad [2.2.24a]$$

where

$$\underline{E}_n = \left[\underline{e}_1^n \ \ \underline{e}_2^n \cdots \underline{e}_n^n\right] \qquad [2.2.24b]$$

$$\underline{E}_m = \left[\underline{e}_1^m \ \ \underline{e}_2^m \cdots \underline{e}_m^m\right] \qquad [2.2.24c]$$

The vectors $\underline{e}_i^j$ are the (jx1) standard Euclidean basis vectors where all the elements in the vector are zero except the i^{th} element which is unity. In equation 2.2.24a, the vectors $\underline{x}$ and $\underline{y}$ are linear combinations of m and n vectors $\underline{x}_i = \underline{e}_i^m\cdot x_i$ and $\underline{y}_i = \underline{e}_i^n\cdot y_i$, respectively. Therefore, in equation 2.2.24a, $\underline{x}$ and $\underline{y}$ are expressed in terms of the standard Euclidean basis vectors and are related by the matrix $\underline{A}$ representing the linear transformation A. Alternatively, $\underline{x}$ and $\underline{y}$ can be decomposed into spatially orthogonal components expressed in terms of basis sets defined by the right and left singular vectors of $\underline{A}$, respectively. If $\underline{A} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T$, equation 2.1.8 can be rewritten as

$$\underline{y} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T\cdot\underline{x} \qquad [2.2.25]$$

Rearranging equation 2.2.25 gives

$$\underline{Z}^T\underline{y} = \underline{\Sigma}\cdot\underline{V}^T\cdot\underline{x} \qquad [2.2.26a]$$

$$\underline{y}^* = \underline{\Sigma}\cdot\underline{x}^* \qquad [2.2.26b]$$

$\underline{y}^*$ will be a linear combination of the columns of $\underline{Z}^T$. $\underline{x}^*$ will be a linear combination of the columns of $\underline{V}^T$. $\underline{y}_i^*$ and $\underline{x}_i^*$ are related by $\underline{\Sigma}$. Since $\underline{\Sigma}$ is a diagonal matrix, the orthogonal components are ordered with respect to their magnitudes such that

$$\underline{y}_i^* = \sigma_i\cdot\underline{x}_i^* \qquad [2.2.27]$$

where σ_i are the singular values of the linear transformation $\mathcal{A}$.

The singular vector sets $\underline{V}$ and $\underline{Z}$ are orthonormal bases which span the fundamental subspaces $\mathrm{Im}(\underline{A}^{-1})$ and $\mathrm{Im}(\underline{A})$, respectively. If

$$\|\underline{x}\|_2 = 1.0 \qquad [2.2.28]$$

the fundamental subspace $\mathrm{Im}(\underline{A})$ is the surface of the ellipsoid representing the output space (the set of all vectors $\underline{y}$ for which equation 2.1.8 is satisfied). The n orthogonal axis of the ellipsoid are aligned with the columns of $\underline{Z}$ ($\underline{z}_i$) and have varying lengths $\{\sigma_i(\underline{A}):1\le i\le n\}$ with the minor and major axes of lengths σ_1 and σ_n, respectively. The spectral norm of the vector $\underline{y}$ will equal the largest singular value of $\underline{A}$ since $\underline{Z}$ and $\underline{V}$ are unitary matrices:

$$\|\underline{Z}^T\cdot\underline{y}\|_2 = \|\underline{\Sigma}\cdot\underline{V}^T\cdot\underline{x}\|_2 \qquad [2.2.29a]$$

$$\|\underline{y}\|_2 = \|\underline{\Sigma}\cdot\underline{x}\|_2 = \|\underline{\Sigma}\|_2 \qquad [2.2.29b]$$

If

$$\underline{y} = 1.0 \qquad [2.2.30]$$

the subspace $\mathrm{Im}(\underline{A}^{-1})$ is the surface of the ellipsoid representing the input or solution space (the set of all vectors $\underline{x}$ for which equation 2.1.8 is satisfied). The m orthogonal axis are aligned with the columns of $\underline{V}$ ($\underline{v}_i$) with axis lengths $(\sigma_i(\underline{A}^{-1}):1\le i\le m)$.

The orthonormal bases $\underline{V} = [\underline{v}_1\ \underline{v}_2\ \dots\ \underline{v}_m]$ and $\underline{Z} = [\underline{z}_1\ \underline{z}_2\ \dots\ \underline{z}_n]$ are formed by rotating the corresponding standard Euclidean bases, $\underline{E}_m = [\underline{e}^m_1\ \underline{e}^m_2 \dots\ \underline{e}^m_m]$ and $\underline{E}_n = [\underline{e}^n_1\ \underline{e}^n_2\ \dots\ \underline{e}^n_n]$, respectively. For $\mathbb{R}^2$, Figure 2.1 shows the rotation angles of ϕ and θ for $\underline{V}$ and $\underline{Z}$, respectively. The change of basis can be used to graphically interpret the linear transformations defined in equation 2.2.24 and 2.2.25 in terms of different basis sets for the subspaces $\mathbb{R}^n$ and $\mathbb{R}^m$. The vector $\underline{x}$, where $\underline{x} = 1.0$, will lie on the spherical subspace defined by the standard Euclidean basis vectors. Through a change of coordinates, $\underline{V}^T$, the vector $\underline{x}$ will be rotated giving $\underline{x}^*$ whose elements are linear combinations of the elements in $\underline{x}$. $\underline{x}^*$ lies on the ellipsoidal subspace representing the transformation in the $\underline{V}$ coordinate system where

$$\left\|\underline{x}^*\right\|_2 = \left\|\underline{V}^T\underline{x}\right\|_2 \le \left\|\underline{V}^T\right\|_2 \cdot \left\|\underline{x}\right\|_2 = 1.0 \qquad [2.2.31]$$

$\underline{x}^*$ is scaled by $\underline{\Sigma}$ forming $\underline{y}^*$. $\underline{y}^*$ lies on the ellipsoid, representing the transformation in the $\underline{Z}$ coordinate system. $\underline{y}^*$ is rotated through another change of coordinates, $\underline{Z}$, forming $\underline{y}$ expressed in terms of the standard Euclidean basis $\underline{E}$. Therefore, each point on the spherical subspace representing the set of vectors $\underline{x}$ ($\left\|\underline{x}\right\|_2 = 1.0$) in standard Euclidean coordinates corresponds to a point on the ellipsoidal subspace representing the set of vectors $\underline{y}$ expressed in standard Euclidean coordinates (Morari, 1989).

Example 2.5: For a linear transformation, $A\colon\mathbb{R}^2 \Rightarrow \mathbb{R}^2$, where

$$\underline{y} = \underline{A} \cdot \underline{x}$$

and

$$\underline{A} = \begin{bmatrix} 1.260 & -0.129 \\ 0.642 & 0.860 \end{bmatrix} \qquad [2.2.32a]$$

the components of the SVD of $\underline{A}$ are:

$$\underline{Z} = \begin{bmatrix} 0.809 & -0.588 \\ 0.588 & 0.809 \end{bmatrix} \qquad [2.2.32b]$$

$$\underline{\Sigma} = \begin{bmatrix} 1.453 & 0 \\ 0 & 0.803 \end{bmatrix} \qquad [2.2.32c]$$

$$\underline{V} = \begin{bmatrix} 0.961 & -0.276 \\ 0.276 & 0.961 \end{bmatrix} \qquad [2.2.32d]$$

If

$$\underline{x} = 1.0 \qquad [2.2.33]$$

the transformation A will map the sphere representing the space of vectors $\underline{x}$ into an ellipsoid representing the set of vectors $\underline{y}$. Figure 2.3 illustrates the propagation of

$$\underline{x} = \underline{e}_1 = [1 \ 0]^T \qquad [2.2.34]$$

through the linear transformation. Figure 2.4 illustrates the propagation of

$$\underline{x} = \underline{v}_1 \qquad [2.2.35]$$

through the process.

If $\underline{A}$ has zero (or nearly zero) singular values, the columns of $\underline{V}$ and $\underline{Z}$ will span ellipsoidal subspaces which have at least one degenerate axis. In this case, the system is overdetermined implying that there is a redundancy of variables (dependency of columns in $\underline{A}$) or redundancy of equations (dependency of rows in $\underline{A}$). If the condition number of $\underline{A}$, $\gamma(\underline{A})$, is large, the minimum singular value of $\underline{A}$ will be considerably smaller in magnitude than the maximum singular value. Since the lengths of the axis of the

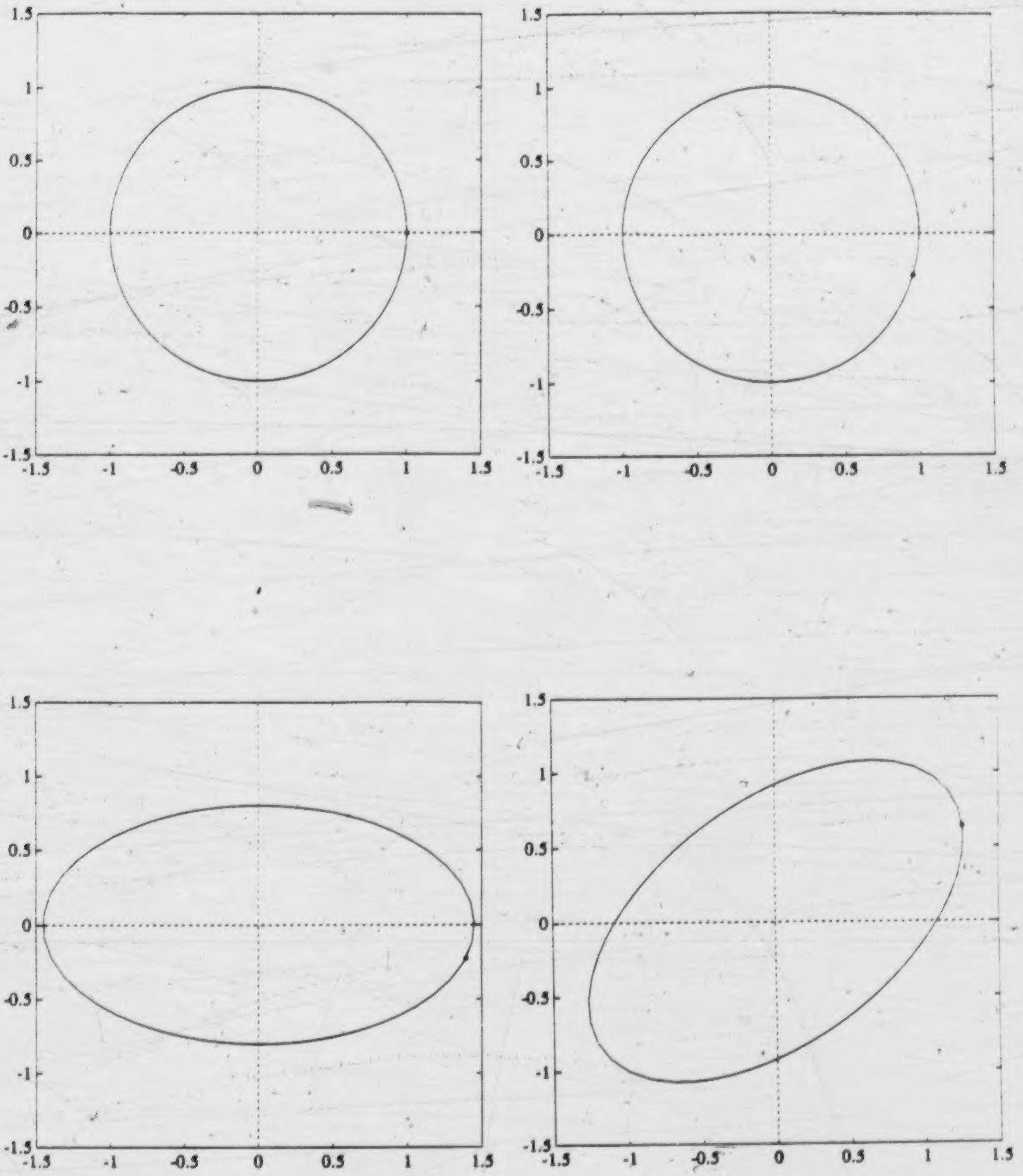

Figure 2.3: Input Signal Propagation Through the Input and Output Spaces of the 2x2 System in Example 2.5 for an Input $\underline{x} = \underline{e}_1$

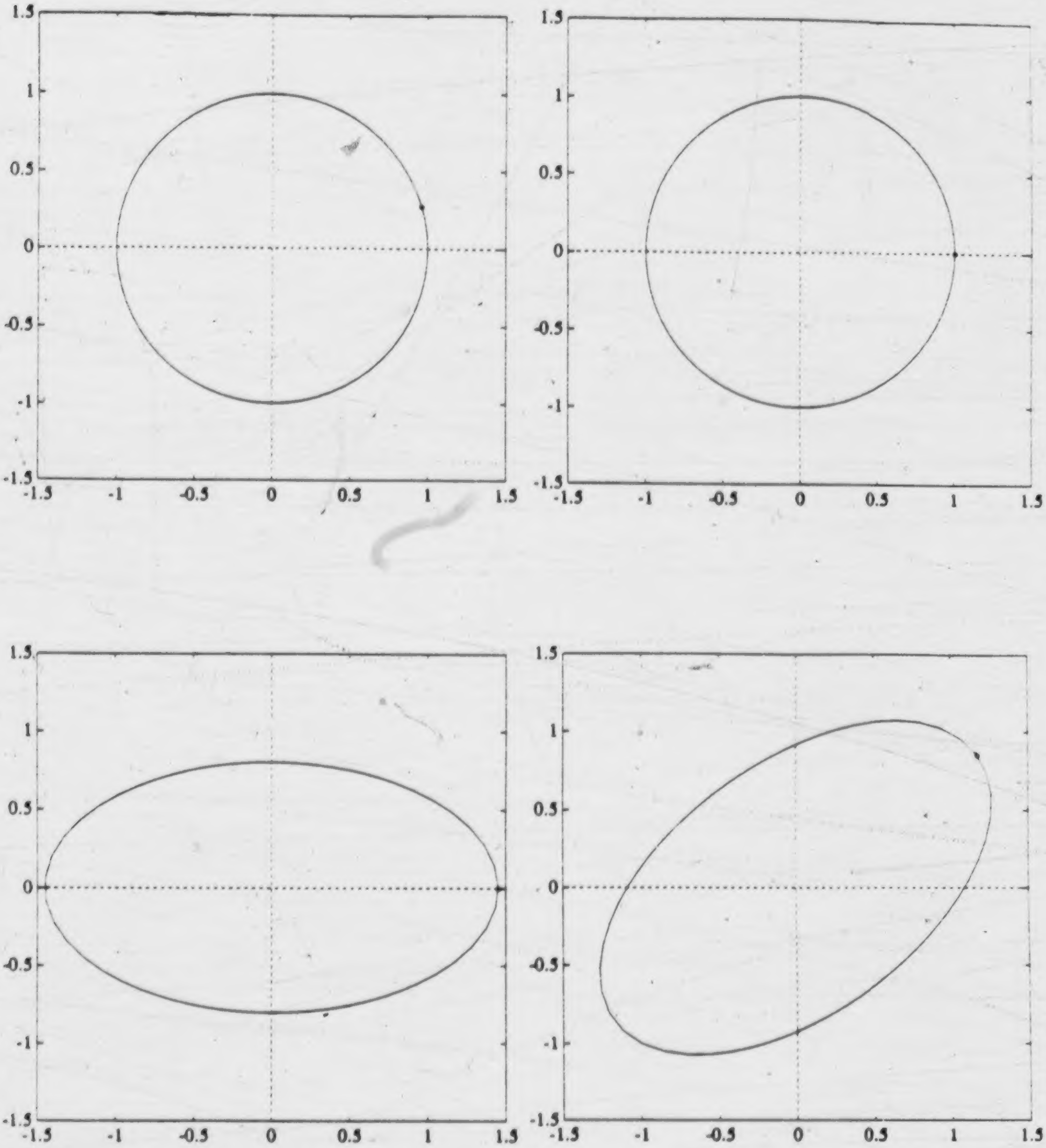

Figure 2.4: **Input Signal Propagation Through the Input and Output Spaces of the 2x2 System in Example 2.5 for an Input $\underline{x} = \underline{v}$**

ellipsoidal subspaces spanned by $\underline{V}$ and $\underline{Z}$ are equal to the singular values of $\underline{A}$, these subspaces will be distorted (Forsythe, 1977).

Example 2.6: The linear transformations $A_1:\mathbb{R}^2 \Rightarrow \mathbb{R}^2$ and $A_2:\mathbb{R}^2 \Rightarrow \mathbb{R}^2$ can be represented by the matrices $\underline{A}_1$ and $\underline{A}_2$ where

$$\underline{A}_1 = \begin{bmatrix} 20 & 0 \\ 0 & 10 \end{bmatrix} \qquad \underline{A}_2 = \begin{bmatrix} 20 & 0 \\ 0 & 1 \end{bmatrix} \qquad [2.2.36]$$

In both cases, the axis directions of the fundamental subspaces $\mathrm{Im}(\underline{A}_i)$ and $\mathrm{Im}(\underline{A}_i^{-1})$ are given by

$$\underline{Z} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \underline{V} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad [2.2.37]$$

However, the axis lengths are different with

$$\underline{\Sigma}_{\underline{A}_1} = \mathrm{diag}(\ 20,\ 10) \qquad [2.2.38a]$$

and

$$\underline{\Sigma}_{\underline{A}_2} = \mathrm{diag}(\ 20,\ 1) \qquad [2.2.38b]$$

Therefore, the condition numbers of the subspaces for $\underline{A}_1$ and $\underline{A}_2$ are 2 and 20, respectively. For $\underline{A}_1$, the subspaces are ellipsoids (spheres) of radius 1. However, $\underline{A}_2$ has a large condition number and is represented by a highly distorted ellipsoid as shown in Figure 2.5. The solution to the linear system, in the direction of the degenerate axis, will be highly suspect due to the lack of information.

Example 2.7: For the linear systems described by the matrices,

$$\underline{A} = \begin{bmatrix} 0.871 & -1.320 \\ 1.578 & -0.095 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 12.071 & -7.377 \\ 12.424 & -6.765 \end{bmatrix} \qquad [2.2.39]$$

the singular vectors are given in equation 2.1.18. The singular values are

$$\underline{\Sigma}_{\underline{A}} = \begin{bmatrix} 2.0 & 0 \\ 0 & 1.0 \end{bmatrix} \qquad \underline{\Sigma}_{\underline{B}} = \begin{bmatrix} 20 & 0 \\ 0 & 0.5 \end{bmatrix} \qquad [2.2.40]$$

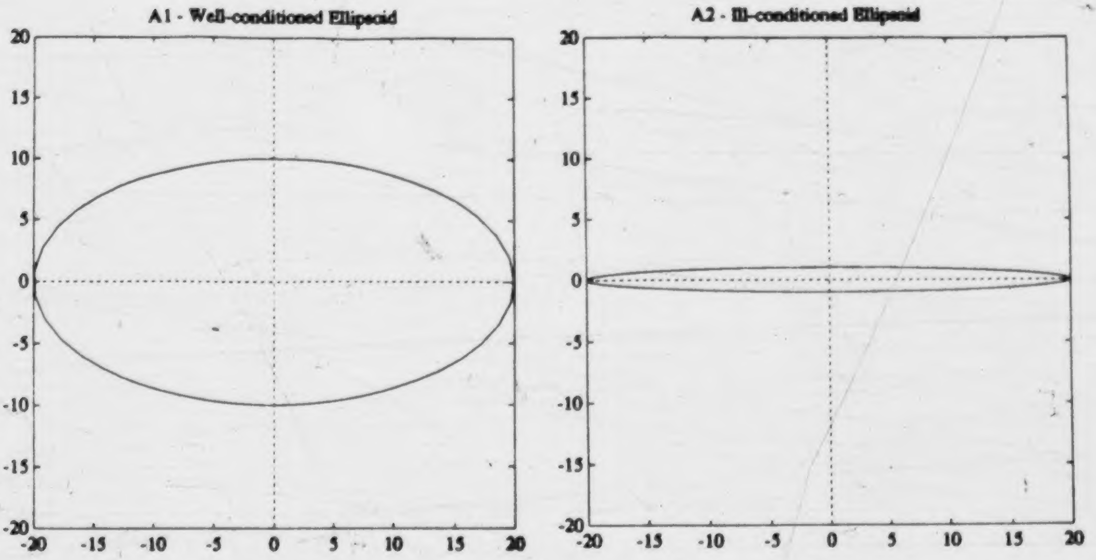

Figure 2.5: The Input and Output Subspaces for the Systems Given in Example 2.6.

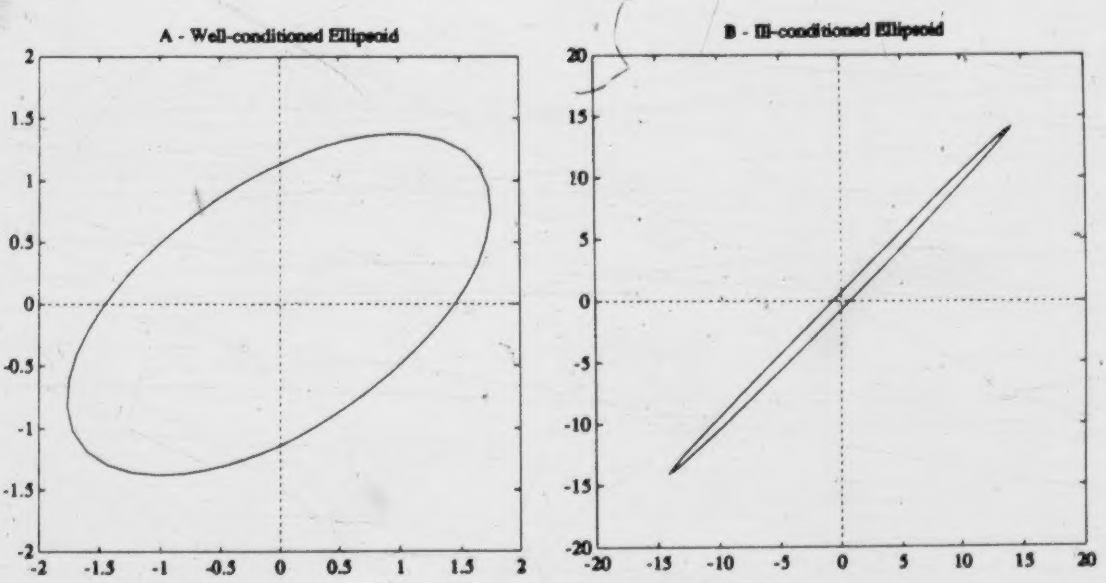

Figure 2.6: The Input and Output Subspaces for the Systems Given in Example 2.6.

The condition number of $\underline{A}$ is 2.0 and the condition number of $\underline{B}$ is 40.0. It can be seen from Figure 2.6 that the ellipsoid representing $\mathrm{Im}(\underline{B})$ is more distorted than the ellipsoid representing $\mathrm{Im}(\underline{A})$. Therefore, the condition number can be pictorially represented by the distortion of an ellipse.

2.2.5 Angle Between Subspaces

The angle between orthonormal subspaces can be used to analyze process interactions in MIMO systems (Klema, 1980). Let $\mathrm{Im}(\underline{A}) = R$ and $\mathrm{Im}(\underline{B}) = S$ define two subspaces of X where $d(X) = n$, $d(R) = r$, $d(S) = s$ and $r+s \leq n$ where $r \leq s$. The angles of inclination between the subspaces R and S are a measure of the separation between the subspaces. The cosines squared of the r angles of inclination between R and S are given by the r largest eigenvalues (or singular values) of $\underline{M} = \underline{A}^T\underline{B}$ (Klema, 1980). If $\underline{M} = \underline{Z} \cdot \underline{\Sigma} \cdot \underline{V}^T$, then

$$\underline{Z}^T \cdot \underline{A}^T \cdot \underline{B} \cdot \underline{V} = \Sigma \tag*{[2.2.41]}$$

and

$$(\underline{A} \cdot \underline{Z})^T \cdot \underline{B} \cdot \underline{V} = \underline{\Sigma} \tag*{[2.2.42]}$$

The matrices representing the subspaces R and S are expressed in terms of the singular vector basis for $\underline{M}$. If $\sigma_i(\underline{M}) = \cos(\Theta_i)$ for $1 \leq i \leq r$, the i^{th} column of $\underline{AZ}$ and the i^{th} column of $\underline{BV}$ will represent a pair of vectors in the subspaces R and S, respectively, with angle Θ_i between them. Θ_i will be the i^{th} inclination angle between the orthogonal basis set of R and S (Klema, 1980). If at least one inclination angle is nonzero, the two subspaces R and S will intersect. If all the inclination angles are 90°, the subspaces will be linearly independent or orthogonal. However, if (at least) one angle is small, the two subspaces will be "close together".

2.3. SVD Algorithms

2.3.1 Eigenvalue-Eigenvector Routines

Conventional diagonalization algorithms used for solving an eigenvalue-eigenvector problem can be used to determine the singular value decomposition of a matrix $\underline{A}$. The eigenvalues and eigenvectors of $\underline{A}^T\underline{A}$ or $\underline{A}\underline{A}^T$ are equivalent to the singular values and left or right singular vectors of $\underline{A}$, respectively. However, the explicit formulation of $\underline{A}^T\underline{A}$ or $\underline{A}\underline{A}^T$ can result in unnecessary numerical inaccuracy due to the increased resolution required in the matrix squaring process (Golub, 1970). If

$$\underline{A} = \begin{bmatrix} 1 & 1 \\ \beta & 0 \\ 0 & \beta \end{bmatrix} \qquad [2.3.1]$$

Squaring the matrix gives

$$\underline{A}^T\underline{A} = \begin{bmatrix} 1+\beta^2 & 1 \\ 1 & 1+\beta \end{bmatrix} \qquad [2.3.2]$$

If β=0.0001, the eigenvalues of $\underline{A}^T\underline{A}$ are λ_1 = 2.0000 and λ_2 = 0.0000 implying the matrix does not have full rank. However, the singular values obtained directly from the SVD algorithm of Golub (1970) are σ_1 = 1.4142 and σ_2 = 0.0001 indicating a full rank matrix. Therefore, if $\underline{A}$ is poorly conditioned, the singular values obtained from the square of $\underline{A}$ will contain numerical errors which will affect the computation of the singular values and singular vectors

2.3.2 Golub and Reinsch (1970) SVD Algorithm

Golub (1970) developed a very efficient and reliable algorithm for calculating the singular value decomposition of a general matrix $\underline{A} \in \mathbb{R}^{nxm}$. Their algorithm iteratively calculates the triplets (σ_i,$\underline{z}_i$, $\underline{v}_i$, i=1,...,n) simultaneously by a two step procedure. Householder transformations are used to reduce the matrix $\underline{A}$ to a bidiagonal form. An iterative procedure involving QR decomposition is used to reduce the super diagonal elements to a negligible size (based on ϵ which is the machine precision) leaving the diagonal matrix $\underline{\Sigma}$ ($\underline{Z}^T\underline{A}\underline{V}=\underline{\Sigma}$). The computed singular vectors are orthonormal within a tolerance of ϵ even with repeated singular values (Moore, 1981).

The algorithm does not perform matrix multiplication but operates on the individual rows and columns of the matrices which decreases the computational load. The majority of the computer load is utilized in the first step since the rate of convergence of the iterations in the second step of the computation, the QR decomposition, is relatively fast (Golub, 1970). The algorithm is stable if (Klema, 1980)

$$\frac{\underline{A} - \underline{Z}\cdot\underline{\Sigma}\cdot V^T}{\underline{A}\cdot\sqrt{m\cdot n}} \simeq \epsilon \qquad [2.3.3]$$

where m and n are the dimensions of $\underline{A}$ and ϵ is the machine precision.

2.3.3 Adaptive Singular Value Decomposition (ASVD) Algorithm

Vandewaale (1984) developed an adaptive singular value decomposition (ASVD) algorithm whereby the algorithm iteratively converges on each triplet (σ_i, $\underline{z}_i$, $\underline{v}_i$) one after the other recursively. This algorithm invokes matrix-vector multiplication, orthogonalization (Gram-Schmidt) and normalization. To improve the execution speed of the algorithm, Vanderwaale suggests performing the orthogonalization procedure every t steps through the iteration instead of each step since this step can be very time consuming. If this is done, accuracy may be sacrificed depending on the choice of t. To decrease the number of iterations required, Vanderwaale suggests using the convergence of the p-1th triplet as a good estimate of the pth triplet. If $\sigma_{p+1}/\sigma_p \simeq 1.0$, the convergence rate of the ASVD algorithm on the pthtriplet can be slow . For this case, Vanderwaale modified the algorithm to speed up the convergence by iterating the pth and the p+1th triplet together. This can be done since only one of the left (resp. right) singular vectors is unique if $\sigma_p = \sigma_{p+1}$. Further iteration is not required since the triplets can then be solved analytically.

The recursive nature of the ASVD algorithm allows the use of the singular vectors and values from the previous time step as the initial guess for their respective counterparts at the next time step. This can be beneficial for on-line real time applications where the matrices vary slowly with time. The ASVD algorithm would conceptually require less computational time than algorithms such as Golub's if only a few triplets need to be calculated. For example, if there is noise in the system, only the triplets (σ_i, z_i, v_i) for which σ_i is greater than some noise criterion δ need to be calculated. Also, if the matrix of interest is structured, there is storage efficiency associated with the ASVD routine because the matrix can be stored by the individual components in a vector (ie. h(1),...h(N)). There is a program asvd.m which was written in MATLAB given in Appendix C which uses Vanderwaale's algorithm to compute the SVD of a matrix. This would be useful in determining the number of terms required in the dynamic matrix for on-line dynamic matrix control calculations.

2.4 Summary

SVD is a reliable method of calculating the rank and condition number of matrices. SVD is also a valuable tool for solving linear equations describing linear transformations since it is effective in dealing with data errors and roundoff errors and can detect linear dependencies. Singular vectors and singular values provide a consistent geometric interpretation of the subspace for the linear transformation since the singular vectors always form an orthonormal basis set for the transformation.

Although SVD solution methods may require more computational time and computer storage than other methods, there are algorithms available which are efficient and numerically sound. For on-line calculations where only a few triplets need to be calculated, the ASVD algorithm may save CPU time and storage space. However, in general, Golub's algorithm is the most efficient method of computing a high precision SVD of a full rank matrix. MATLAB utilizes the LINPACK routine for SVD which is equivalent to Golub's routine.

Chapter 3 Principal Component Analysis (PCA) of Linear Dynamic Systems

3.0 Introduction

Principal component analysis (PCA) is a technique for analyzing vector signals in a dynamic linear system using the computational machinery of SVD. Moore (1979a, 1981) used principal component analysis of a state space process model to analyze the controllability and observability of asymptotically stable continuous linear systems. Several researchers (Lau, 1985, Moore, 1981, Maurath, 1985, Callaghan, 1986, Clarke, 1987) have examined the use of PCA in the analysis and design of predictive controllers and MIMO control strategies. PCA has also been used to linearize continuous time non-linear systems using a least squares approximation criteria (Moore, 1979b).

In this chapter, several different forms of process models are described. In particular, it is shown how the shape of the state space for a particular process can be changed via a linear transformation. The application of principal component analysis to the matrix representing the linear transformation between an input sequence and an output sequence for a time invariant dynamic system is explored. The results are then extended to describe the controllable and observable subspaces of the system.

3.1 Linear Dynamic Systems Models

There are two general types of process models which represent the dynamic and steady state behaviour of linear systems: parametric and non-parametric. Non-parametric models are expressed as transient functions or frequency functions of the process response (Isermann, 1981). Parametric models express the relationship between a process input and an output as a linear equation. Parametric models explicitly contain parameters in the form of a transfer function, weighting sequence or state space model. Time series analysis of measured process signals or examination of process responses provide methods of identifying the structure of the process model. The parameters can be estimated from process response data either recursively or non-recursively using least squares techniques. Model based control applications utilize parametric models because they describe the process in a manner which is easy to manipulate mathematically.

3.1.1 Discrete Transfer Function Process Models

3.1.1.1 Auto regressive Moving Average (ARMA) Process Models

The most common parametric process representation is an auto regressive moving average (ARMA) model:

$$y(t) = \underset{\substack{\text{process}\\ \text{model}}}{\frac{z^{-d}B(z^{-1})}{A(z^{-1})}}\, u(t) + \underset{\substack{\text{noise}\\ \text{model}}}{\frac{C(z^{-1})}{D(z^{-1})}}\, \xi(t) = G(z^{-1})\cdot u(t) + N(z^{-1})\cdot \xi(t) \qquad [3.1.1]$$

where $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$ and $D(z^{-1})$ are polynomials in the backward shift operator, z^{-1}, and d is the process delay. The present output, y(t), depends on past inputs (moving average) and outputs (auto regressive) as well as past noise components, $\xi(t)$. For MIMO systems, $\underline{G}(z^{-1})$ is a transfer function matrix where the i^{th} input is related to the j^{th} output via the transfer function $G_{ji}(z^{-1})$ such that

$$\underline{y}(t) = \underline{G}(z^{-1})\cdot\underline{u}(t) + \underline{N}(z^{-1})\cdot\underline{\xi}(t) \qquad [3.1.2]$$

In this representation, the transfer function relating the i^{th} input to the j^{th} output is independent of the transfer function relating the noise to the output. The transfer functions are explicitly defined in terms of the backward shift operator.

Several variations on the basic model type can be obtained depending on the choice of the polynomials (A,B,C,D). For example, if the effects of noise are neglected ($C(z^{-1})=0$), a DARMA (deterministic auto regressive moving average) model is obtained:

$$y(t) = \frac{z^{-d}B(z^{-1})}{A(z^{-1})}\,u(t) = G(z^{-1})\cdot u(t) \qquad [3.1.3]$$

which relates the output response to the past inputs and outputs only. The transfer functions relating the output to the input and the noise can be expressed in terms of a common denominator as

$$y(t) = \frac{z^{-d}\cdot B^{*}(z^{-1})}{A^{*}(z^{-1})}\,u(t) + \frac{C^{*}(z^{-1})}{A^{*}(z^{-1})}\,\xi(t) = G(z^{-1})\cdot u(t) + N(z^{-1})\cdot\xi(t) \qquad [3.1.4a]$$

where

$$A^{*}(z^{-1}) = A(z^{-1})\cdot D(z^{-1}) \qquad [3.1.4b]$$

$$B^{*}(z^{-1}) = B(z^{-1})\cdot D(z^{-1}) \qquad [3.1.4c]$$

$$C^{*}(z^{-1}) = A(z^{-1})\cdot C(z^{-1}) \qquad [3.1.4d]$$

One model of this type is a CARIMA (controlled auto regressive integrated moving average) model used by Mohtadi (1986) in the implementation of GPC (generalized predictive control):

$$y(t) = \frac{z^{-d}B(z^{-1})}{A(z^{-1})} \cdot u(t) + \frac{C(z^{-1})}{A(z^{-1})} \cdot \frac{\xi(t)}{\Delta} \qquad [3.1.5]$$

where Δ is a differencing operator, and $\xi(t)$ is an independent, random signal. Kailath (1980) and Isermann (1981) provide a good overview of the ARMA model and several variations.

The ARMA transfer function models require a finite number of parameters. If the orders of $A(z^{-1})$ and $B(z^{-1})$ are 'n' and 'm', respectively, with d representing any time delay, then (m+n+d) or fewer parameters must be identified for the deterministic process model. The orders of the polynomials $C(z^{-1})$ and $D(z^{-1})$ determine the number of parameters required in the noise model. To identify an ARMA model from process data, an estimation of the process model orders, m and n, and the time delay are required apriori. There are identification methods available for determining the noise parameters such as extended least squares (ELS) or time series techniques but an estimate of the noise model structure is required apriori. In the case of the CARIMA or similar models where the noise and input/output models are expressed in terms of a common denominator, the identification of the noise and I/O models is not independent. If the actual order of the process is unknown, iterative searches using correlation analysis can be used to estimate the order. If the assumed and actual process orders are different, the resultant ARMA model may not represent the true behaviour of the process. In practice, a low order model, usually first or second order, is used to describe the process and the control system is designed to deal with the plant-model mismatch by detuning the controllers.

3.1.1.2 Impulse Response (Weighting Sequence) and Step Response Process Models

Another type of process representation is a weighting sequence model which utilizes the impulse response of the system to explicitly determine the effects of past inputs on the present output. In the impulse response model, the present output is expressed as a linear combination of all the past inputs and noise terms weighted by their respective impulse response parameters. The process model given in equation 3.1.1 can be written as

$$y(t)=\sum_{i=1}^{\infty} h_i z^{-i}u(t) + \left[1+\sum_{i=1}^{\infty} f_i z^{-i}\right]\xi(t) = H(z^{-1})u(t) + F(z^{-1})\xi(t) \qquad [3.1.6]$$

where $\{h_i : i = 1,2,..,\infty\}$ are the impulse response parameters and $H(z^{-1})$ and $F(z^{-1})$ are polynomials of infinite degree in the backward shift operator. The effects of the inputs on the process outputs can be determined directly from the impulse response model since the present output does not depend explicitly on the past outputs. The magnitude of the i^{th} impulse parameter, h_i, reflects how significant the $(t-i)^{th}$ past input, $u(t-i)$, is on the present output, $y(t)$. In a MIMO system,

$$\underline{y}(t)=\sum_{i=1}^{\infty} \underline{h}_i z^{-i}\underline{u}(t) + \left[\underline{1}+\sum_{i=1}^{\infty} \underline{f}_i z^{-i}\right]\underline{\xi}(t) = \underline{H}(z^{-1})\underline{u}(t) + \underline{F}(z^{-1})\underline{\xi}(t) \qquad [3.1.7]$$

In the deterministic case, the impulse response parameters can be obtained directly from the impulse response of the process as shown in Figure 3.1. If the process signal/noise ratio is high, determination of the parameters directly from the process response to a pulse (or step input) may be very inaccurate. In this case, other techniques such as cross-correlation analysis with a PRBS input can be used to obtain a more accurate process model. Also, impulse response parameters can be obtained using recursive or batch least squares techniques. However, the vectors (or matrices) involved in the identification can be quite large depending on the number of parameters in the weighting sequence.

Similarly, a step response model of the process can be determined from the process response to a step input as shown in Figure 3.2. In this representation, the present output is a linear combination of the past incremental moves of the manipulated variables and the noise. The process model is given by

$$y(t) = S(z^{-1})\Delta u(t) + F(z^{-1})\xi(t) = \sum_{i=1}^{\infty} s_i z^{-i}\Delta u(t) + \left[1+\sum_{i=1}^{\infty} f_i z^{-i}\right]\xi(t) \quad [3.1.8]$$

where $\{s_i;\ i=1,2,..,\infty\}$ are the step response parameters. Since the impulse response of a process is the derivative of the step response, the impulse response parameters, h_i, are equal to the difference between consecutive step response parameters. Thus, the i^{th} impulse parameter is given by

$$h_i = s_i - s_{i-1} = \Delta s_i \quad [3.1.9]$$

In a MIMO system, the step response model is given by

$$\underline{y}(t) = \underline{S}(z^{-1})\Delta\underline{u}(t) + \underline{F}(z^{-1})\underline{\xi}(t) = \sum_{i=1}^{\infty} \underline{s}_i z^{-i}\Delta\underline{u}(t) + \left[\underline{I}+\sum_{i=1}^{\infty} \underline{f}_i z^{-i}\right]\underline{\xi}(t) \quad [3.1.10]$$

where the i^{th} impulse response vector is given by

$$\underline{h}_i = \underline{s}_i - \underline{s}_{i-1} = \Delta\underline{s}_i \quad [3.1.11]$$

Theoretically, a weighting sequence model requires an infinite number of parameters to completely model the process. However, in practical applications, it is assumed that the infinite sequence can be truncated after N sampling intervals where τ (= N*Ts) represents the settling time or dominant time constant of the process. The resulting (deterministic) finite

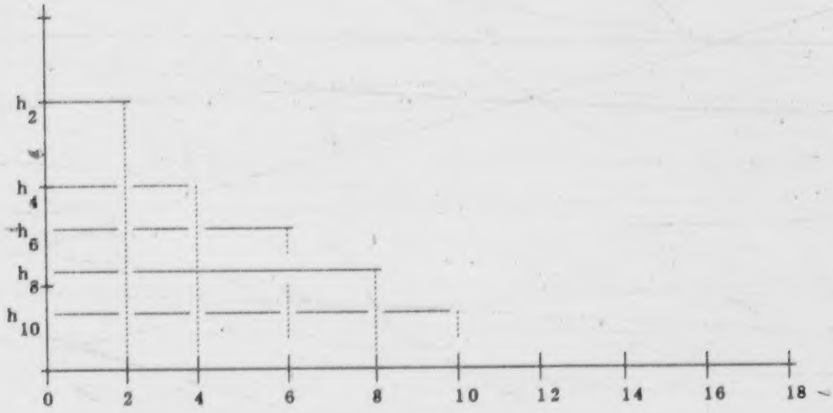

Figure 3.1: Impulse Response of a System

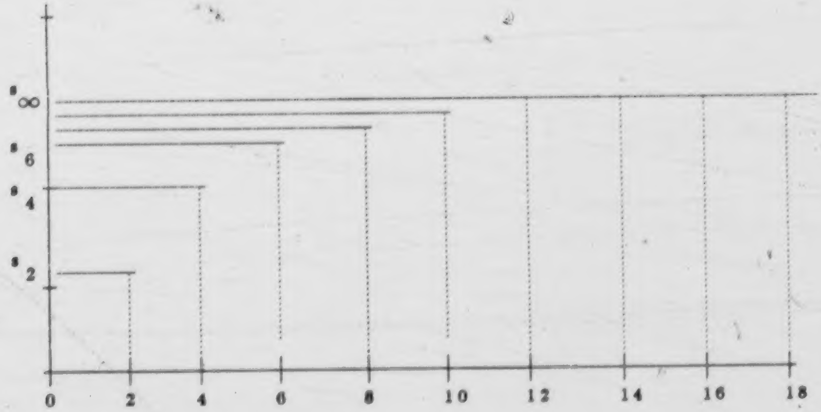

Figure 3.2: Step Response of a System

impulse response model (FIR) is given by

$$y(t) = \sum_{i=1}^{N} h_i z^{-i} u(t) + \sum_{i=N+1}^{\infty} h_i z^{-i} u(t)$$

$$= H(z^{-1})u(t) + E \qquad [3.1.12]$$

where E represents the truncation error. If $h_i \simeq 0$ for i>N, the truncation error or model error should be insignificant since for any stable plant, $h_i \Rightarrow 0$ as $i \Rightarrow \infty$. The rate at which the parameters approach zero depends on the dominant time constant of the process. For a system with a fast response time (small dominant time constant), the system responds quickly to changes in the input signal and is influenced significantly by the most recent inputs. For a sluggish system (large dominant time constant), the system does not respond as quickly to input changes. Therefore, the effects of the most recent inputs will not be as significant. Fewer impulse response parameters will be required to adequately model a fast process as compared to a sluggish system.

Similarly, the finite step response model is given by

$$y(t) = \sum_{i=1}^{N} s_i \Delta u(t-i) + \sum_{i=N+1}^{\infty} s_i \Delta u(t-i)$$

$$= S(z^{-1})\Delta u(t) + y_o \qquad [3.1.13]$$

In this case, the truncation of the infinite series produces a "steady state" term, y_o. Assuming that the process was initially at steady state, y_o can be estimated as

$$y_o = \sum_{i=N+1}^{\infty} s_i z^{-i} \Delta u(t) \simeq s_{\infty} u(t-N-1) \qquad [3.1.14]$$

where $s_{N+1} = s_{N+2} = .. = s_{\infty}$ is the steady state gain of the process and u(t-N-1) is the cumulative effect of the past incremental control actions

$\{\Delta u(t-i): i = N+1,..,\infty\}$ (Navratril, 1988).

Identification of a weighting sequence process model does not require an apriori knowledge of the process order. However, a truncation error is introduced to avoid identification of an infinite number of parameters. The accuracy of these FIR models can be improved by increasing the length of the weighting sequence since modeling errors arise from the truncation of the infinite series. An upper bound on this truncation error can be determined which can indicate the robustness of the resulting model based control scheme. In contrast, an apriori knowledge of process order is required for identification of a transfer function or an ARMA model. The determination of an upper bound on the unmodeled plant dynamics is not straightforward with an ARMA model.

For underdamped systems with persistent oscillations, a weighting sequence model will require a very large number of parameters to accurately model the process. Any model based control scheme which used these models would require a comparatively large amount of CPU processing time. Unstable processes can not be modeled accurately with a weighting sequence model since an infinite number of parameters is required. On the other hand, an ARMA model can be used to model unstable or poorly damped processes reasonably well.

Continuous time transfer function and weighting sequence process models can be expressed in terms of Laplace transforms and convolution integrals in a similar manner as the preceding discrete models (eqn's 3.1.1, 3.1.6, 3.1.8):

$$y(s) = \underset{\text{process model}}{\frac{e^{-\tau_d}B(s)}{A(s)}}\,u(s) + \underset{\text{noise model}}{\frac{C(s)}{D(s)}}\,\xi(s) = G(s)\cdot u(s) + N(s)\cdot\xi(s) \qquad [3.1.15]$$

$$y(t) = \int_0^\infty h(\tau - i)\cdot u(\tau)d\tau + \int_0^\infty f(\tau - i)\cdot\xi(\tau)d\tau \qquad [3.1.16]$$

3.1.2 State Space Process Models

A state space model represents the process by a set of differential equations. An internal state structure describes the process dynamics. For a linear, time invariant process with p inputs, m outputs and n internal states, the state space model is given by

$$\dot{\underline{x}}(t) = \underline{A}\underline{x}(t) + \underline{B}\underline{u}(t) \qquad [3.1.17a]$$

$$\underline{y}(t) = \underline{C}^T\underline{x}(t) + \underline{D}\underline{u}(t) \qquad [3.1.17b]$$

for continuous time systems where $\underline{A} \in \mathbb{R}^{nxn}$, $\underline{B} \in \mathbb{R}^{nxp}$, $\underline{C} \in \mathbb{R}^{pxn}$ and $\underline{D} \in \mathbb{R}^{mxp}$. The solution to the state space equation of a time-invariant continuous linear system is

$$\underline{x}(t) = \underline{x}(0)e^{\underline{A}t} + \int_0^t e^{\underline{A}\tau}\underline{B}\underline{u}(\tau)d\tau \qquad [3.1.18]$$

The state space system can be expressed in discrete time as

$$\underline{x}(k+1) = \underline{\underline{A}}\underline{x}(\kappa) + \underline{\underline{B}}\underline{u}(k) \qquad [3.1.19a]$$

$$\underline{y}(k) = \underline{\underline{C}}^T\underline{x}(k) + \underline{\underline{D}}\underline{u}(k) \qquad [3.1.19b]$$

where

$$\underline{\underline{A}} = e^{\underline{A}t} \qquad [3.1.19c]$$

and

$$\underline{\underline{B}} = \int_0^t e^{\underline{A}(t-\tau)}\underline{B}(\tau)d\tau \qquad [3.1.19d]$$

For each state space model, an unique transfer function model can be defined as in equation 3.1.15 and 3.1.1 where

$$\underline{G}(s) = \underline{C}^{T}(s\underline{I}-\underline{A})^{-1}\underline{B} \qquad [3.1.20a]$$

$$\underline{G}(z) = \underline{C}^{T}(z\underline{I}-\underline{A})^{-1}\underline{B} \qquad [3.1.20b]$$

for continuous time and discrete time systems, respectively. Similarly, a weighting sequence model can be determined from a state space model of the process. If the process has one input and one output, the weighting sequence model corresponding to the state space model given by equation 3.1.17 is

$$y(k) = \sum_{i=1}^{\infty} h_i z^{-i} u(k) = \sum_{i=1}^{\infty} \underline{c}\underline{A}^{i-1}\underline{b}u(k) = \underline{c}(z\underline{I}-\underline{A})^{-1}\underline{b}u(k) \qquad [3.1.21]$$

If the input signal to the process is an impulse (or pulse), the impulse parameters are given by (discrete time)

$$h(k) = \mathcal{Z}^{-1}H(z) = \underline{c}\underline{A}^{k-1}\underline{b} \quad k\geq 1 \qquad [3.1.22]$$

where $t = kT_s$ where T_s is the sampling time and $\mathcal{Z}^{-1}$ is the inverse z-transform. Similarly, in continuous time, the impulse response parameters can be shown to be given by

$$h(t) = \mathcal{L}^{-1}H(s) = \underline{c}e^{\underline{A}t}\underline{b} \qquad [3.1.23]$$

where $\mathcal{L}^{-1}$ represents the inverse Laplace transform of H(s). In a multivariable system,

$$\underline{h}(k) = \underline{CA}^{k-1}\underline{B} \qquad \text{or} \qquad \underline{h}(t) = \underline{C}e^{\underline{A}t}\underline{B} \qquad [3.1.24]$$

where $\underline{h}(k)$ (or $\underline{h}(t)$) is an pxm matrix.

For a given input-output model represented by a transfer function (matrix) or weighting sequence, there are several equivalent state space models which will generate the same output response to a given input sequence. Although each of the state space models exhibit the same input and output properties, the internal states are scaled differently. These equivalent state space models are related through linear transformations of the state variables. Let $(\underline{A},\underline{B},\underline{C})$ represent a minimal realization of the continuous system given in equations 3.1.17a and 3.1.17b. If $\underline{T}$ is a matrix representing a linear transformation which scales the internal states such that

$$\underline{\bar{x}}(t) = \underline{T}\underline{x}(t) \qquad [3.1.25]$$

the scaled system becomes

$$\dot{\underline{\bar{x}}}(t) = \underline{\bar{A}}\underline{\bar{x}}(t) + \underline{\bar{B}}\underline{u}(t) \qquad [3.1.26a]$$

$$\underline{y}(t) = \underline{\bar{C}}\underline{\bar{x}}(t) + \underline{\bar{D}}\underline{u}(t) \qquad [3.1.26b]$$

where

$$\underline{\bar{A}} = \underline{T}^{-1}\underline{A}\underline{T} \qquad [3.1.26c]$$

$$\underline{\bar{B}} = \underline{T}^{-1}\underline{B} \qquad [3.1.26d]$$

$$\underline{\bar{C}} = \underline{C}\underline{T} \qquad [3.1.26e]$$

Example 3.1: There are several state space models which will describe the input-output characteristics of the following SISO system described by the continuous transfer function model (Moore, 1981)

$$G(s) = \frac{s^2 + 15s + 50}{s^4 + 5s^3 + 33s^2 + 79s + 50} \qquad [3.1.27]$$

A continuous state space model, in controller canonical form, which produces the same input-output behaviour is

$$\underline{A} = \begin{bmatrix} -5 & -33 & -79 & -50 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \underline{C} = \begin{bmatrix} 0 & 1 & 15 & 50 \end{bmatrix} \qquad [3.1.28]$$

The state variables in the model given in equation 3.1.28 can be transformed via a linear transformation to a Jordan canonical form. Let the transformation matrix be equal to

$$\underline{T} = \begin{bmatrix} 0.005+0.007i & 0.005-0.007i & -0.125 & 1.0 \\ -0.037+0.016i & -0.037-0.016i & 0.250 & -1.0 \\ -0.040-0.196i & -0.400+0.196i & -0.500 & 1.0 \\ 1.0 & 1.0 & 1.0 & -1.0 \end{bmatrix} \qquad [3.1.29]$$

where the columns of $\underline{T}$ are equal to the eigenvectors of $\underline{A}$. The transformed (continuous) state space model is given by

$$\underline{\tilde{A}} = \begin{bmatrix} -1+4.90i & 0 & 0 & 0 \\ 0 & -1-4.90i & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad \underline{\tilde{B}} = \begin{bmatrix} 0.361+0.376i \\ 0.361-0.376i \\ 0.320 \\ 0.042 \end{bmatrix} \quad \underline{\tilde{C}} = \begin{bmatrix} -0.365+0.368i \\ -0.365-0.368i \\ -3.00 \\ 36.0 \end{bmatrix}$$

[3.1.30]

where the diagonal elements of $\underline{\tilde{A}}$ are the eigenvalues of $\underline{A}$. The state space model $\{\underline{\tilde{A}},\underline{\tilde{B}},\underline{\tilde{C}}\}$ in equation 3.1.30 is equivalent to the state space model $\{\underline{A},\underline{B},\underline{C}\}$ given in equation 3.1.28 and will produce the same input/output behaviour. The corresponding impulse response and step response parameters can be obtained from equation 3.1.27 (or equation 3.1.28) or directly from the impulse response and step response as shown in Figure 3.3.

For the discrete time system described by the function model

$$G(z) = \frac{0.065z^3 + 0.025z^2 - 0.071z + 0.020}{z^4 - 3.320z^3 + 4.312z^2 - 2.594z + 0.607} \qquad [3.1.31]$$

A corresponding discrete space model in controller canonical form is

$$\underline{A} = \begin{bmatrix} 3.320 & -4.312 & 2.594 & -0.607 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \underline{C}^T = \begin{bmatrix} 0.065 \\ 0.025 \\ -0.071 \\ 0.020 \end{bmatrix}$$

[3.1.32]

The corresponding Jordan canonical state space model is given by

$$\underline{\tilde{A}} = \begin{bmatrix} 0.80+0.43i & 0 & 0 & 0 \\ 0 & 0.80-0.43i & 0 & 0 \\ 0 & 0 & 0.82 & 0 \\ 0 & 0 & 0 & 0.91 \end{bmatrix} \quad \underline{\tilde{B}} = \begin{bmatrix} 0.03+0.04i \\ 0.03-0.04i \\ 0.03 \\ 0.004 \end{bmatrix} \quad \underline{\tilde{C}}^T = \begin{bmatrix} -0.37+0.37i \\ -0.37-0.37i \\ -3.00 \\ 36.0 \end{bmatrix}$$

[3.1.33]

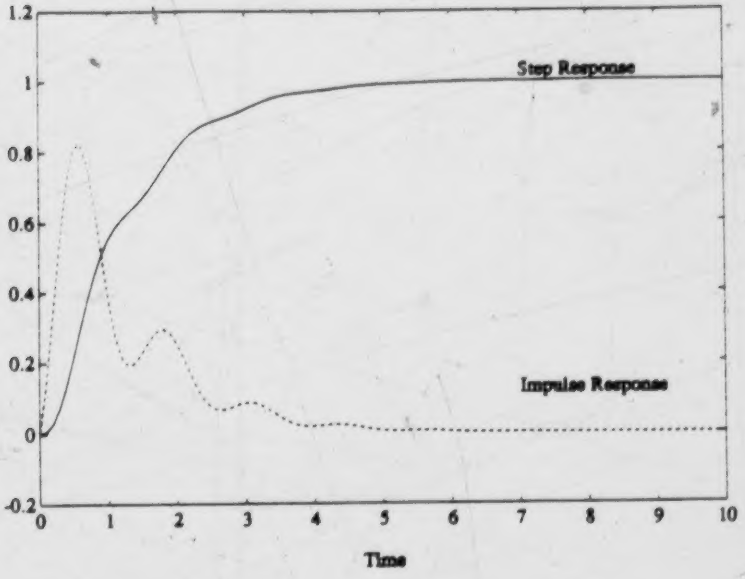

Figure 3.3a: Unit Impulse Response and Unit Step Response for the Continuous System Given in Equation 3.1.27

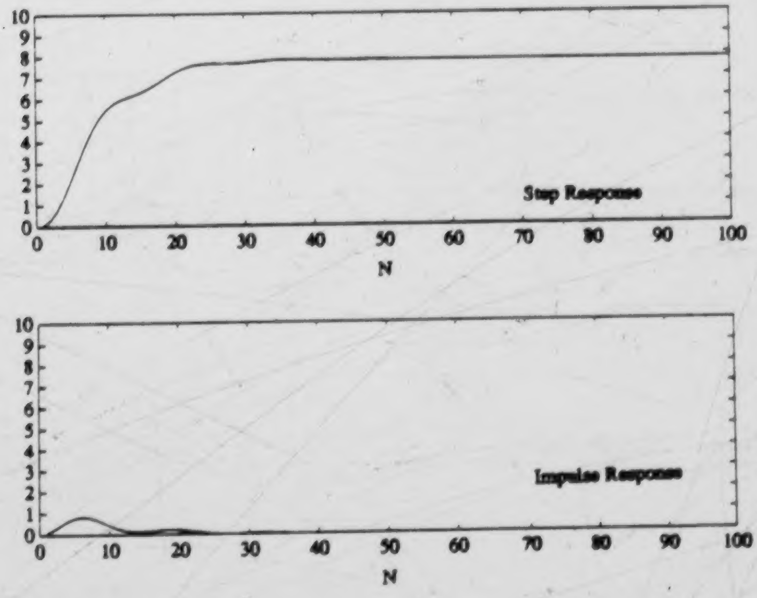

Figure 3.3b: Unit Impulse and Unit Step Response for the Discrete System Given in Equation 3.1.31

The state space models given in equations 3.1.32 and 3.1.33 will describe the same input/output behaviour as the transfer function model in equation 3.1.31 as shown in Figure 3.3b.

3.2 PCA of Linear Time Invariant Dynamic Systems

3.2.1 Continuous Time Systems

PCA has been used in statistics to analyze vector time signals. The same analysis techniques can be used to assess the total energy and the distribution of energy in dynamic processes. The energy in the vector signals is an indication of persistency of excitation in the dynamic system. Once a process model has been obtained which represents the process, principal component analysis (PCA) can be used to define the image space and the kernel space of the linear transformation (see chapter 2).

Let $\underline{u}(t) \in \mathbb{R}^p$, $\underline{y}(t) \in \mathbb{R}^m$, and $\underline{G}(t) \in \mathbb{R}^{m\times p}$ represent the linear time invariant system shown in Figure 3.4, where

$$\underline{y}(t) = \int_0^T \underline{G}(t-\tau)\underline{u}(\tau)d\tau \qquad [3.2.1]$$

such that $\underline{u}(\cdot)$ is piecewise continuous over the interval [0,T] (Moore,

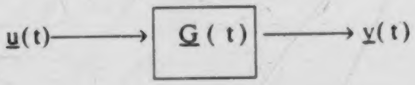

Figure 3.4: Linear Time Invariant System

1981). A positive, symmetric, semi-definite matrix, the Grammian ($\underline{W}$), can be defined as (Moore, 1979a, 1981):

$$\underline{W}^2 = \int_0^T \underline{G}(t)\cdot\underline{G}(t)^T dt = \underline{Z}\cdot\underline{\Sigma}^2\cdot\underline{Z}^T \qquad [3.2.2a]$$

or

$$\underline{W}^{2^T} = \int_0^T \underline{G}(t)^T\cdot\underline{G}(t) dt = \underline{Z}^T\cdot\underline{\Sigma}^2\cdot\underline{Z} \qquad [3.2.2b]$$

where $\Sigma^2=\mathrm{diag}(\sigma_1^2,\sigma_2^2,\ldots,\sigma_m^2)$. The singular values of $\underline{W}^2$, σ_i, are real, non-negative numbers ordered such that,

$$\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_m^2 \geq 0 \qquad [3.2.3]$$

The matrix has mutually orthogonal singular vectors,

$$\{\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_m\} \qquad [3.2.4]$$

which are the columns of $\underline{Z}$. Since the Grammian, $\underline{W}^2$, is a symmetric matrix, the singular values and vectors of $\underline{W}^2$ are the same as the eigenvalues and eigenvectors of $\underline{W}^2$.

The singular vectors of $\underline{W}^2$ span $\mathbb{R}^m$ forming an orthonormal basis for the subspace. $\underline{G}(t)$ can be expressed in terms of this orthonormal basis as (Moore, 1981)

$$\underline{G}(t) = \underline{z}_1\underline{g}_1^T(t) + \underline{z}_2\underline{g}_2^T(t) + \ldots + \underline{z}_m\underline{g}_m^T(t) \qquad [3.2.5]$$

where $\underline{g}_i^T(t) = \underline{z}_i^T\underline{G}(t)$ for $1 \leq i \leq m$. The i^{th} principal component of $\underline{G}(t)$ is defined as $\underline{z}_i\underline{g}_i^T$ where $\underline{z}_i$ is the i^{th} component vector of $\underline{G}(t)$ and $\underline{g}_i^T$ is the i^{th} component function vector of $\underline{G}(t)$. The magnitude of the i^{th} principal component of $\underline{G}(t)$, σ_i, represents the length of the corresponding component function vector, $\underline{g}_i$. The sum of the component magnitudes of $\underline{G}(t)$ over the interval [0,T] represents the total energy in the signal set (Moore, 1981)

$$\int_0^T \underline{G}(t)\ {}^2_F\ dt = \sum_{i=1}^{m} \sigma_i^2 = \text{trace}\left\{\int_0^T \underline{G}(t)\cdot\underline{G}(t)^T dt\right\} \qquad [3.2.6]$$

where $\cdot\ _F$ is the Frobenius norm. The component vectors, $\underline{z}_i$, and component magnitudes, σ_i, define an ellipsoidal region of the subspace $\mathbb{R}^m$ which reflects the spatial distribution of the energy in the system. Assuming that $\sigma_m \geq 0$, the subspace containing the set of vectors $\underline{z} \in \mathbb{R}^m$ for which

$$\int_0^T \underline{z}^T\underline{G}(t)\ {}^2\ dt = \int_0^T \underline{g}^T(t)\ {}^2 dt = 1 \qquad [3.2.7]$$

is an ellipsoid with semi-axes of length σ_i and direction $\underline{z}_i$ for $1 \leq i \leq m$.

The image of the convolution map given in equation 3.2.1 is the ellipsoidal subspace of $\mathbb{R}^m$ defined by the component magnitudes and component vectors of $\underline{G}(t)$ over [0,T] (Moore, 1981). A set of vectors

$$\overset{*}{\underline{y}} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{p} \qquad [3.2.8]$$

can be defined where $\underline{p} = 1.0$. The subspace which contains all vectors $\overset{*}{\underline{y}}$ which satisfy equation 3.2.8 is an ellipsoid with semi-axis of length σ_i and direction $\underline{z}_i$. If an input signal

$$\underline{u}(t) = \underline{G}^T(T-t)\cdot\underline{a} \qquad [3.2.9a]$$

is applied to the system described by equation 3.2.8 where

$$\underline{a} = \underline{Z}\cdot\underline{\Sigma}^{-1}\cdot\underline{p} \qquad [3.2.9b]$$

$\underline{y}(t)$ will be driven to $\overset{*}{\underline{y}}$ at time T (Moore, 1981). If

$$\overset{*}{\underline{y}} = \sigma_i\cdot\underline{z}_i \qquad [3.2.10]$$

and $\underline{u}(\cdot)$ satisfies the norm bound

$$\int_0^T \underline{u}(t)\ {}^2 dt = \int_0^T \underline{u}(t)\cdot\underline{u}^T(t)dt = 1 \qquad [3.2.11]$$

the minimum norm input function will be the i^{th} component function vector normalized and reflected in time (Moore, 1981)

$$\underline{u}(t) = \underline{G}^T(T-t)\underline{z}_i\sigma_i^{-1} = \underline{f}_i(T-t)/\sigma_i \qquad [3.2.12]$$

The principal components can be determined based on sampled data. If $\underline{G}(t)$ is sampled at 0, t_1, t_2,...,t_N, a rectangular approximation of the integral can be used to calculate the Grammian, $\underline{W}^2$. The interval [0,T] is divided into N evenly spaced sample points. For a small enough sample time and a large enough sample size N (Moore, 1981)

$$\underline{W}^2 = \int_0^T \underline{G}(t)\underline{G}(t)^T dt \cong \frac{T-0}{N}\sum_{i=1}^{N} \underline{G}(t_i)\underline{G}(t_i)^T = \underline{DD}^T \qquad [3.2.13a]$$

where

$$\underline{D} = \left[\frac{T-0}{\sqrt{N}}\right]^{1/2} \cdot \left[\underline{G}(0)\ \underline{G}(t_1) \dots \underline{G}(t_N)\right] \qquad [3.2.13b]$$

Therefore, the component magnitudes and vectors of $\underline{G}(t)$ over the interval [0,T] are approximated by the singular values and vectors of $\underline{D}$. The vector sequence $\underline{D}$ can be decomposed into spatially orthogonal vectors, $\underline{z}_i$, ordered with respect to their component magnitudes $\sigma_i(\underline{D})$ where $\underline{W}^2 = \underline{DD}^T = \underline{Z}\cdot\underline{\Sigma}^2\cdot\underline{Z}^T$ and $\underline{\Sigma} = \mathrm{diag}(\sigma_1,\sigma_2,..,\sigma_m)$.

Example 3.2: For the system represented by the matrix

$$\underline{G}(t) = \begin{bmatrix} e^{-t} & 0.005e^{-0.1t} \\ 0.5e^{-0.5t} & e^{-t} \end{bmatrix} \qquad [3.2.14]$$

the Grammian can be evaluated over the interval [0,T] using equations 3.2.13a and 3.2.13b to approximate the integral. If the sampling time is 0.5, the value of the Grammian evaluated over the interval [0,5] is

$$\underline{W}^2 = \begin{bmatrix} 2.502 & 1.520 \\ 1.520 & 3.502 \end{bmatrix} \qquad [3.2.15a]$$

where

$$\underline{Z}_W = \underline{V}_W = \begin{bmatrix} 0.586 & -0.810 \\ 0.810 & 0.586 \end{bmatrix} \qquad \underline{\Sigma}_W^2 = \begin{bmatrix} 4.602 & 0 \\ 0 & 1.402 \end{bmatrix} \qquad [3.2.15b]$$

The component vectors of $\underline{G}(t)$ are given by the columns of $\underline{Z}_W$ and the component magnitudes are given by the diagonal elements of $\underline{\Sigma}_W$. At t=0, the principal components of $\underline{G}(0)$ will be given by

$$g_1^T(0) = z_1^T\underline{G}(0) = \begin{bmatrix} 0.991 & 0.810 \end{bmatrix} \qquad [3.2.16a]$$

$$g^T(0) = z^T\underline{G}(0) = \begin{bmatrix} -0.517 & 0.582 \end{bmatrix} \qquad [3.2.16b]$$

and

$$\begin{aligned} \underline{G}(0) &= z_1 g_1^T + z_2 g_2^T \\ &= \begin{bmatrix} 0.581 & 0.475 \\ 0.803 & 0.656 \end{bmatrix} + \begin{bmatrix} 0.419 & -0.471 \\ -0.303 & 0.341 \end{bmatrix} \\ &= \begin{bmatrix} 1.00 & 0.005 \\ 0.50 & 1.0 \end{bmatrix} \qquad [3.2.17] \end{aligned}$$

$g_1^T(0)$ and $g_2^T(0)$ are orthogonal components representing each vector signal in $\underline{G}(t)$ (each column of $\underline{G}(t)$.

As the time interval increases, the Grammian changes resulting in a change in the component magnitudes or singular values. However, the component vectors of the Grammian do not change. Therefore, the ellipsoid representing the spatial distribution of energy in the vector signals simply expands along its principal axes. There is an increase in the energy of the signal set represented by $\underline{G}(t)$ as indicated by the increase in the component magnitudes. For the interval [0,20]

$$\underline{W}^2 = \begin{bmatrix} 5.004 & 3.034 \\ 3.034 & 7.012 \end{bmatrix} \qquad [3.2.18a]$$

where

$$\underline{Z}_W = \underline{V}_W = \begin{bmatrix} 0.586 & -0.811 \\ 0.811 & 0.586 \end{bmatrix} \qquad \underline{\Sigma}_W^2 = \begin{bmatrix} 9.204 & 0 \\ 0 & 2.282 \end{bmatrix} \qquad [3.2.18b]$$

For the interval [0,40], the Grammian is

$$\underline{W}^2 = \begin{bmatrix} 7.076 & 4.291 \\ 4.291 & 9.916 \end{bmatrix} \qquad [3.2.19a]$$

where

$$\underline{Z}_W = \underline{V}_W = \begin{bmatrix} 0.586 & -0.811 \\ 0.811 & 0.586 \end{bmatrix} \qquad \underline{\Sigma}_W^2 = \begin{bmatrix} 13.016 & 0 \\ 0 & 3.977 \end{bmatrix} \qquad [3.2.19b]$$

Figure 3.5 shows the ellipsoids representing the image space of $\underline{G}$ and the distribution of energy at different time intervals.

If $\underline{G}(t)$ is rank deficient (rank = r<m), there will be columns (and rows) of $\underline{G}(t)$ which are linearly dependent over the interval [0,T]. The linearly dependent columns correspond to zero component magnitudes of $\underline{W}^2$ which represent degenerate axis of the ellipsoidal region corresponding to the image space of the system. Principal components can be used to approximate $\underline{G}(t)$, in the least squares sense, with a system of lower rank $\underline{G}_r(t)$ over the interval [0,T]. Over the class of piecewise continuous functions $\underline{G}(t)$ with dimension r ($1 \le r < m$), the function

$$\underline{G}_r(t) = \sum_{i=1}^{r} \underline{z}_i \underline{f}_i^T(t) \qquad [3.2.20]$$

can be defined which will minimize the Frobenius norm and the spectral norm of the approximation error (Moore, 1981). The approximation error will be given by

$$E_F = \int_0^T \quad \underline{G}(t) - \underline{G}_r(t)\ _F^2 \quad dt = \sum_{i=r+1}^{m} \sigma_i^2 \qquad [3.2.21a]$$

$$E_S = \max_{\underline{z} \neq 1} \int_0^T \underline{z}^T (\underline{G}(t) - \underline{G}_r(t))\ _2^2 \ dt = \sigma_{r+1}^2 \qquad [3.2.21b]$$

The singular values and the singular vectors of $\underline{D}$ in equation 3.2.7b,

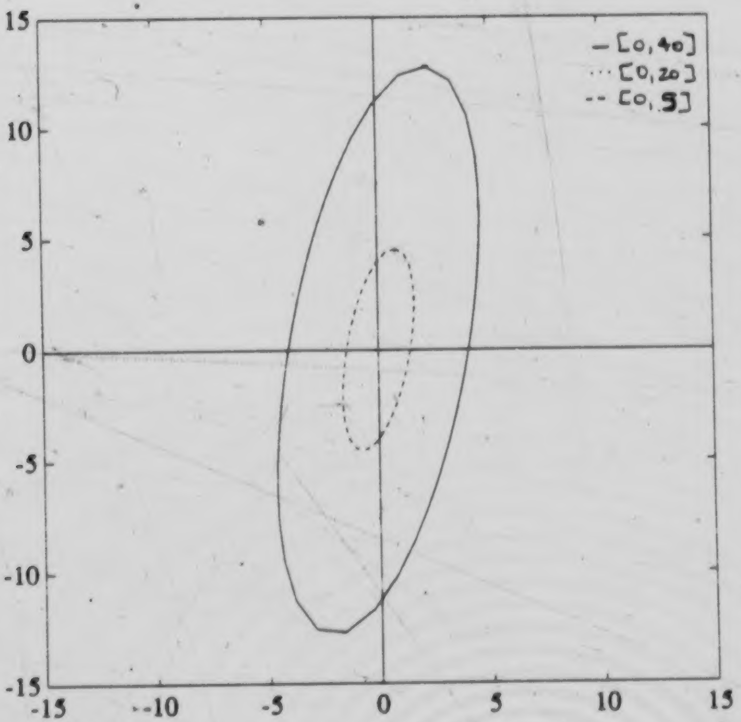

Figure 3.5: Subspaces Representing the Distribution of Energy in the System (the Image of $\underline{G}$) for the System in Example 3.2

can be used to find $\underline{G}_r(t)$. The component magnitudes or singular values of $\underline{D}$ reflect the closeness of the approximations of various dimension r. However, for the continuous time case, $\underline{G}_r(t)$, given by equation 3.2.20, will minimize the approximation error only at the sample points (Moore, 1979a).

As discussed in section 2.2.1 for the linear static equation

$$\underline{y} = \underline{A} \cdot \underline{x}$$

a perturbation (or error) in the data vector $\underline{y}$ will result in a error in the solution vector $\underline{x}$. The condition number of $\underline{A}$ indicates the worst case magnification of the error in the data vector into the error of the solution vector. In dynamic systems, the condition number of the Grammian $\underline{W}^2$ indicates the sensitivity of the system to perturbations in $\underline{u}$ and $\underline{y}$ (Moore, 1979a). For the system given in equation 3.2.1, Moore (1979a) showed that a perturbation in $\underline{u}$, $\Delta\underline{u}$, will be magnified by $\gamma(\underline{W})$in the output vector $\underline{y}$ giving $\underline{y}+\Delta\underline{y}$

$$\left[\frac{\int_0^T \underline{G}(t)\Delta\underline{u}^{\,2} dt}{\int_0^T \underline{G}(t)\underline{u}^{\,2} dt}\right]^{1/2} = \left[\frac{\sigma^2_{\max}(\underline{W})}{\sigma^2_{\min}(\underline{W})}\right]^{1/2} \frac{\Delta\underline{u}}{\underline{u}} = \gamma(\underline{W}) \cdot \frac{\Delta\underline{u}}{\underline{u}} \qquad [3.2.22]$$

If $\underline{u}_y(t)$ is the minimum norm solution of the system, Moore (1979a) showed that a perturbation in $\underline{y}$, $\Delta\underline{y}$, will be magnified by $\gamma(\underline{W}^T)$ in the minimum norm solution to the system where $\Delta\underline{u}(t) = \underline{u}_{y+\Delta y}(t) - \underline{u}_y(t)$

$$\left[\frac{\int_0^{t_1} \Delta\underline{u}(t)^{\,2} dt}{\int_0^{t_1} \underline{u}_y(t)^{\,2} dt}\right]^{1/2} = \left[\frac{\sigma^2_{\max}(\underline{W}^T)}{\sigma^2_{\min}(\underline{W}^T)}\right]^{1/2} \frac{\Delta\underline{y}}{\underline{y}} = \gamma(\underline{W}^T) \cdot \frac{\Delta\underline{y}}{\underline{y}} \qquad [3.2.23]$$

Perturbations in the matrix $\underline{G}(t)$ ($\Delta\underline{G}(t)$) will alter the component magnitudes and vectors of $\underline{G}(t)$. Let the perturbed system

$$\underline{G}_{\Delta}(t) = \underline{G}(t) + \Delta\underline{G}(t) \qquad [3.2.24]$$

be represented by the Grammian $\underline{W}^2_{\Delta}$ given by

$$\underline{W}^2_{\Delta} = \int_0^T \underline{G}_{\Delta}(t)\underline{G}^T_{\Delta}(t)dt \qquad [3.2.25]$$

with component magnitudes $\sigma_{i\Delta}$ where

$$\sigma^2_{1\Delta} \geq \sigma^2_{2\Delta} \geq \ldots \geq \sigma^2_{n\Delta} \geq 0 \qquad [3.2.26]$$

A perturbation in $\underline{G}(t)$ will change the component magnitudes by no more than the component magnitudes of the perturbation such that (Moore, 1979a)

$$\left| \sigma^2_i - \sigma^2_{i\Delta} \right| \leq \int_0^T \Delta\underline{G}(t)\Delta\underline{G}(t)^T dt \qquad [3.2.27]$$

A similar result was shown for the linear system $\underline{y} = \underline{A}\cdot\underline{x}$ in equation 2.2.5.

3.3 Controllability and Observability Analysis

3.3.1 Continuous Time Systems

PCA can be used to analyze the controllability and observability of the system represented by the state space equations

$$\dot{\underline{x}}(t) = \underline{A}\cdot\underline{x}(t) + \underline{B}\cdot\underline{u}(t) \qquad [3.3.1a]$$

$$\underline{y}(t) = \underline{C}\cdot\underline{x}(t) + \underline{D}\cdot\underline{u}(t) \qquad [3.3.1b]$$

where $\underline{x}(t) \in \mathbb{R}^n$, $\underline{u}(t) \in \mathbb{R}^p$, and $\underline{y}(t) \in \mathbb{R}^m$. Within the state space structure there are four internal subspaces: the controllable, uncontrollable, observable and unobservable subspaces. A system is state controllable at any time if given an initial state $\underline{x}(t_o)$, there is a piecewise continuous $\underline{u}(t)$ which will take all states from their initial states, $\underline{x}(t_o)$, to any final state $\underline{x}(t)$ in a finite time. Therefore, the controllable subspace is the smallest subspace which contains the state response to every piecewise continuous input signal $\underline{u}(t)$ for the system with initial states $\underline{x}(0) = \underline{0}$. The uncontrollable subspace is orthogonal to the controllable subspace. A system is state observable if a knowledge of the input $u(\tau)$ and the output $y(\tau)$ over a finite time segment $0<\tau\leq t$ completely determines $x(t)$. Therefore, the unobservable subspace is the largest subspace for which arbitrary piecewise continuous input signals $\underline{u}(t)$ can be applied to the system with no output response. The observable subspace is orthogonal to the unobservable subspace. These subspaces can be related to the image and kernel spaces of a controllability and an observability Grammian.

If $\underline{G}(t)$ represents the impulse response matrix of the system

$$\underline{y}(t) = \underline{G}(t)\cdot\underline{u}(t) \qquad [3.3.2]$$

then

$$\underline{G}(t) = \underline{C}e^{\underline{A}t}\underline{B} \qquad [3.3.3]$$

A controllability Grammian is defined as (Moore, 1979a, 1981)

$$\underline{W}_C^2 = \int_0^T e^{\underline{A}\tau}\underline{B}\,\underline{B}^T e^{\underline{A}^T\tau}\,d\tau = \underline{Z}_c\underline{\Sigma}_c\underline{Z}_c^T \qquad [3.3.4a]$$

where

$$\underline{\Sigma}_c^2 = \mathrm{diag}\{\sigma_{ci}^2\},\quad \sigma_{c1}^2 \geq \sigma_{c2}^2 \geq \ldots \geq \sigma_{cn}^2 \geq 0 \qquad [3.3.4b]$$

An observability Grammian is defined as (Moore, 1979a, 1981)

$$\underline{W}_o^2 = \int_0^T e^{\underline{A}^T\tau}\underline{C}^T\underline{C}e^{\underline{A}\tau}\,d\tau = \underline{V}_o\underline{\Sigma}_o\underline{V}_o^T \qquad [3.3.5a]$$

where

$$\underline{\Sigma}_o = \mathrm{diag}\,\{\sigma_{oi}\},\quad \sigma_{o1}^2 \geq \sigma_{o2}^2 \geq \ldots \geq \sigma_{on}^2 \geq 0 \qquad [3.3.5b]$$

These unique Grammians, $\underline{W}_c^2$ and $\underline{W}_o^2$, are positive, semi-definite, symmetric matrices which satisfy the Lyapunov equations

$$\underline{A}\underline{W}_c^2 + \underline{W}_c^2\underline{A}^T = -\underline{B}\underline{B}^T \qquad [3.3.6a]$$

$$\underline{A}^T\underline{W}_o^2 + \underline{W}_o^2\underline{A} = -\underline{C}^T\underline{C} \qquad [3.3.6b]$$

$\underline{W}_c^2$ and $\underline{W}_o^2$ are obtained from the solutions of these Lyapunov equations.

For the system shown in Figure 3.6, the controllable subspace can be defined by monitoring the state responses $\{\underline{x}_i(t)\}$ with initial states of zero, $\{\underline{x}(0)=\underline{0}\}$, to the sequence of impulse test signals

$$\underline{u}_i(t) = e_i^p \cdot \delta(t) \qquad 1 \leq i \leq p \qquad [3.3.7]$$

where $\delta(t)$ is the scalar unit impulse function, e_i^p is the i^{th} column of the mxm identity matrix and d(t)=0 (Moore, 1981). From equation 3.1.18, the state response is given by

$$\underline{X}(t) = e^{\underline{A}t}\underline{B} \qquad [3.3.8]$$

where $\underline{X}(t)=[\underline{x}_1(t)\ \ \underline{x}_2(t)\ldots\underline{x}_p(t)] \in \mathbb{R}^{nxp}$. Therefore, the image space of $\underline{X}(t)$ with the least dimension for all $t \in [0,T]$ will define the controllable

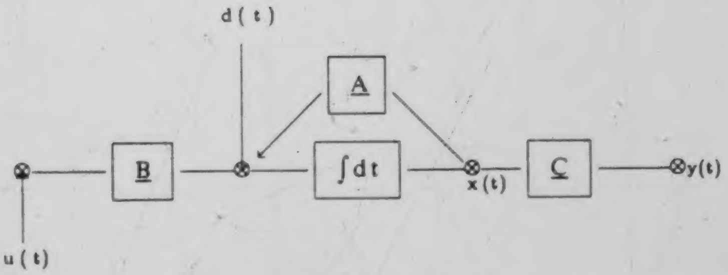

Figure 3.6: State Responses to Test Signals (u(t) and d(t)) (Moore, 1981)

subspace. Since

$$\underline{W}_c^2 = \int_0^T e^{\underline{A}\tau}\underline{B}\underline{B}^T e^{\underline{A}^T\tau}\,d\tau = \int_0^T \underline{X}(\tau)\cdot\underline{X}(\tau)d\tau \qquad [3.3.9]$$

the controllable subspace is also the image space of $\underline{W}_c^2$. Therefore, component vectors of $\underline{W}_c^2$, $\underline{z}_{ci}$, corresponding to non-zero component magnitudes, σ_{ci}, span the controllable subspace and define an ellipsoidal region in the state space with axis of lengths $\sigma_{ci}>0$ and directions $\underline{z}_{ci}$ which can be reached from a zero initial state with input vectors satisfying

$$\int_0^T \underline{u}(t)^2 dt \le 1 \qquad [3.3.10]$$

The uncontrollable subspace is orthogonal to the controllable subspace and corresponds to the kernel space of $\underline{W}_c$. This is the subspace corresponding to zero component magnitudes of $\underline{W}_c$.

Similarly, the observable subspace corresponds to the image space of $\underline{W}_o^2$ (Moore, 1981). For the system shown in Figure 3.6, the observable subspace can be defined by monitoring the output response to a sequence of impulse test signals

$$\underline{d}_i(t) = e_i^n\cdot\delta(t) \quad 1\le i\le n \qquad [3.3.11]$$

where e_i^n is the i^{th} column of the nxn identity matrix and $\underline{u}(t) = 0$ (Moore, 1981). The solution to the state equations is

$$\underline{Y}(t) = \underline{C}\cdot e^{\underline{A}t} \qquad [3.3.12]$$

where $\underline{Y}(t) = [\underline{y}_1(t)\ \underline{y}_2(t)\ \dots\ \underline{y}_n(t)] \in \mathbb{R}^{mxn}$. Therefore, the kernel space of $\underline{Y}(t)$ with the greatest dimension for all $t \in [0,T]$ defines the unobservable subspace. Since

$$\underline{W}_o^2 = \int_0^T e^{\underline{A}^T\tau}\underline{C}^T\underline{C}e^{\underline{A}\tau}\,d\tau = \int_0^T \underline{Y}^T(t)\cdot\underline{Y}(t)dt \qquad [3.3.13]$$

the component vectors of $\underline{W}_o^2$, $\underline{z}_{oi}$, corresponding to zero component magnitudes, σ_{oi}, span the unobservable subspace. The component vectors corresponding to non-zero component magnitudes span the observable subspace and define an ellipsoidal region within the state space with axis lengths σ_{oi} in directions $\underline{z}_{oi}$ which defines the set of all initial conditions which satisfy

$$\int_0^T \underline{C} e^{\underline{A}t} \underline{x}(0) \;^2 \, dt = 1 \qquad [3.3.14]$$

Scaling of the internal states of the state space system given in equation 3.1.17a and 3.1.17b by a linear transformation, $\underline{T}$, will change the singular values and vectors describing the controllability and observability subspaces. Therefore, the "shape" of these subspaces will be affected. If the states are improperly scaled, the ellipsoids will be very distorted or thin indicating that some of the states are nearly uncontrollable and/or unobservable.

If the states are perturbed by $\Delta\underline{x}$ such that an input $\underline{u}_{x\Delta}$ is required to drive the output from $\underline{y}(0)$ to $\underline{y}(T)$, then (Moore, 1979a)

$$\left[\frac{\int_0^T \underline{u}_{x\Delta} - \underline{u}_x \;^2 dt}{\int_0^T \underline{u}_x \;^2 dt} \right]^{1/2} \leq \gamma(\underline{W}_c) \cdot \frac{\Delta\underline{x}}{\underline{x}} \qquad [3.3.15]$$

where $\gamma(\underline{W}_c)$ reflects the conditioning of the system $(\underline{A}, \underline{B}, \underline{C})$ with respect to point wise state control on the interval [0,T]. Similarly, if the initial states of the system are perturbed by $\Delta\underline{x}_0$ which produce an output $\underline{y}_\Delta(t) = \underline{C} e^{\underline{A}t} \underline{x}_{0\Delta}$ over the interval [0,T], then (Moore, 1979a)

$$\left[\frac{\int_0^T \underline{y}(t)-\underline{y}_\Delta(t)\ {}^2dt}{\int_0^T \underline{y}(t)\ {}^2dt}\right]^{1/2} \leq \frac{1}{\gamma(\underline{W}_o)}\cdot\frac{\Delta\underline{x}_0}{\underline{x}_0} \qquad [3.3.16]$$

where $\gamma(\underline{W}_o)$ reflects the conditioning of the system $(\underline{A},\underline{B},\underline{C})$ with respect to zero input state observation. If the controllable and observable subspaces are very distorted, the condition numbers of the corresponding Grammians will be very large. In this case, small perturbations in the states can result in large errors in the predicted process output to a given input sequence.

Example 3.6: For the system described by the continuous state space model in equation 3.1.28, the controllability Grammian (equation 3.3.4) is

$$\underline{W}_c^2 = \begin{bmatrix} 0.213 & 0 & 0.007 & 0 \\ 0 & 0.007 & & 0.0005 \\ -0.007 & 0 & 0.0005 & 0 \\ 0 & -0.0005 & 0 & 0.0002 \end{bmatrix} \qquad [3.3.17a]$$

$$\underline{W}_o^2 = \begin{bmatrix} 0.005 & 0.023 & 0.127 & 0.250 \\ 0.003 & 0.135 & 0.740 & 1.476 \\ 0.127 & 0.740 & 4.336 & 8.878 \\ 0.250 & 1.476 & 8.878 & 18.581 \end{bmatrix} * 10^2 \qquad [3.3.17b]$$

The singular vectors of $\underline{W}_c^2$ and $\underline{W}_o^2$, the columns of $\underline{Z}_c$ and $\underline{Z}_o$ respectively, are

$$\underline{Z}_c = \begin{bmatrix} 0.999 & 0 & 0.034 & 0 \\ 0 & 0.998 & 0 & -0.064 \\ -0.034 & 0 & -0.999 & 0 \\ 0 & -0.064 & 0 & -0.998 \end{bmatrix} \qquad [3.3.18a]$$

$$\underline{Z}_o = \begin{bmatrix} 0.012 & -0.074 & 0.021 & 0.997 \\ 0.072 & -0.368 & -0.927 & -0.009 \\ 0.432 & -0.824 & 0.361 & -0.074 \\ 0.899 & 0.426 & -0.099 & 0.023 \end{bmatrix} \qquad [3.3.18b]$$

The singular vectors represent a orthonormal basis set for the ellipsoidal controllable and observable subspaces in $\mathbb{R}^4$, respectively, for this system. The singular values are

$$\{\sigma_i(\underline{W}_c^2): \ 0.213,\ 0.007,\ 0.0002,\ 0.0001\} \qquad [3.3.19a]$$

and

$$\{\sigma_i(\underline{W}_o^2): \ 2.296{*}10^3,\ 783,\ 0.446,\ 0.0595\} \qquad [3.3.19b]$$

which represent the length of the orthonormal axis of the ellipsoidal subspaces. In this realization, there are at least two states which are almost uncontrollable as indicated by the small singular values of $\underline{W}_c^2$. Also, there are at least two states which are relatively unobservable as indicated by the relative magnitudes of $\sigma_{min}(\underline{W}_o^2)$ and $\sigma_{max}(\underline{W}_o^2)$. The condition numbers for the controllable subspace and the observable subspace are

$$\gamma(\underline{W}_c) = 1.69{*}10^3 \qquad [3.3.20a]$$

$$\gamma(\underline{W}_o) = 3.86{*}10^4 \qquad [3.3.20b]$$

respectively. Although this realization is more controllable than observable, the high condition numbers indicate that there will be states which are almost uncontrollable and unobservable.

Example 3.7: For the continuous state space system

$$\underline{A} = \begin{bmatrix} -0.5183 & 1.4503 & 0.3911 \\ -1.4503 & -2.1954 & -4.7533 \\ 0.3911 & 4.7533 & -0.6297 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} -0.7729 \\ -0.8047 \\ 0.3373 \end{bmatrix} \quad \underline{C}^T = \begin{bmatrix} -0.7729 \\ 0.8047 \\ 0.3373 \end{bmatrix} \qquad [3.3.21]$$

the controllability and observability Grammians are given by

$$\underline{W}_c^2 = \begin{bmatrix} 0.5763 & 0 & 0 \\ 0 & 0.1475 & 0 \\ 0 & 0 & 0.0904 \end{bmatrix} = \underline{W}_o^2 \qquad [3.3.22]$$

where $\underline{Z}_c = \underline{Z}_0 = \underline{I}$ and

$$\{\sigma_i(\underline{W}_c^2) = \sigma_i(\underline{W}_o^2) : 0.5763,\ 0.1475,\ 0.0904\} \qquad [3.3.23]$$

Therefore

$$\gamma(\underline{W}_c) = \gamma(\underline{W}_o) = 6.38 \qquad [3.3.24]$$

In this case, the system is as controllable as it is observable since the condition number of the controllability Grammian and the observability Grammian are equal. Also, this system will be considerably easier to control and observe than the system in example 3.6.

3.3.2 Discrete Time Systems

The controllability and observability subspaces for discrete time systems can be defined in a similar manner as the continuous system. If

$$\underline{G}(k) = \underline{C}\underline{A}^{k-1}\underline{B} \qquad [3.3.25]$$

represents the impulse response matrix of the discrete system given in equations 3.1.19a and 3.1.19b, the discrete time controllability and observability Grammians are defined as (Silverman, 1980)

$$\underline{W}_c^2 = \sum_0^k \underline{A}^i \underline{B}\underline{B}^T (\underline{A}^T)^i = \underline{Z}_c \underline{\Sigma}_c^2 \underline{Z}_c^T \qquad [3.3.26a]$$

$$\underline{W}_o^2 = \sum_0^k (\underline{A}^T)^i \underline{C}^T \underline{C}\underline{A}^i = \underline{V}_o \underline{\Sigma}_o^2 \underline{V}_o^T \qquad [3.3.26b]$$

respectively. These unique Grammians, $\underline{W}_c^2$ and $\underline{W}_o^2$, are positive, semi-definite, symmetric matrices which satisfy the discrete time Lyapunov equations

$$\underline{A}\underline{W}_c^2\underline{A}^T + \underline{W}_c^2 = -\underline{B}\underline{B}^T \qquad [3.3.27a]$$

$$\underline{A}^T\underline{W}_o^2\underline{A} + \underline{W}_o^2 = -\underline{C}^T\underline{C} \qquad [3.3.27b]$$

As the sampling time of a system approaches zero, the discrete time controllability and observability Grammians will approach the corresponding continuous Grammians (Moore, 1979a)

$$(\underline{W}_c^2)_{cont} = \underset{t_s \Rightarrow 0}{\text{limit}} \left\{ \frac{1}{t_s} (\underline{W}_c^2)_{dis} \right\} \qquad [3.3.28a]$$

$$(\underline{W}_o^2)_{cont} = \underset{t_s \Rightarrow 0}{\text{limit}} \left\{ t_s (\underline{W}_c^2)_{dis} \right\} \qquad [3.3.28b]$$

It can be shown that the discrete time Grammians are related to the system controllability and observability matrices, $\underline{P}$ and $\underline{Q}$ respectively, such that (Silverman, 1980)

$$\underline{W}_c = \underline{P}\underline{P}^T \qquad [3.3.29a]$$

$$\underline{W}_o = \underline{Q}^T\underline{Q} \qquad [3.3.29b]$$

where

$$\underline{P} = [\ \underline{B}\ \ \underline{AB}\ \ \underline{A}^2\underline{B} \cdots \underline{A}^{n-1}\underline{B}\ ] \qquad [3.3.30a]$$

$$\underline{Q} = [\ \underline{C}\ \ \underline{CA}\ \ \underline{CA}^2 \cdots \underline{CA}^{n-1}\ ] \qquad [3.3.30b]$$

The determinant of the system controllability and observability matrices can be used to assess controllability and observability of the discrete state space system. The state space model is said to be uncontrollable if det($\underline{P}$)≅0 and is unobservable if det($\underline{Q}$)≅0. Although the determinant of $\underline{P}$ and $\underline{Q}$ indicates whether the system is controllable/observable or uncontrollable/unobservable, it does not indicate the "degree" of controllability or observability. On the other hand, the condition number of $\underline{P}$ and $\underline{Q}$ will indicate whether the matrices are close to rank deficient. If the condition number is large, the determinant of $\underline{P}$ and $\underline{Q}$ will be close to zero indicating that there are some states which may be almost uncontrollable and unobservable, respectively. However, the number of uncontrollable and unobservable states will be unknown. On the other hand, the singular values of $\underline{W}_c^2$ and $\underline{W}_o^2$ indicate uncontrollability/unobservability of the system as well as the number of states which may be unobservable/uncontrollable.

Example 3.8: For the discrete state space system in example 3.1 (equation 3.1.32), the controllability and observability Grammians are given by

$$\underline{W}_c^2 = \begin{bmatrix} 168.50 & 96.44 & 41.76 & 10.56 \\ 96.44 & 57.05 & 25.60 & 6.71 \\ 41.76 & 25.60 & 12.02 & 3.32 \\ 10.56 & 6.71 & 3.32 & 1.00 \end{bmatrix} \tag*{[3.3.31a]}$$

and

$$\underline{W}_o^2 = \begin{bmatrix} 0.670 & -1.324 & 0.952 & -0.247 \\ -1.324 & 2.661 & -1.935 & 0.505 \\ 0.952 & -1.935 & 1.416 & -0.371 \\ -0.247 & 0.505 & -0.371 & 0.097 \end{bmatrix} \tag*{[3.3.31b]}$$

The singular values are

$$\sigma_i(\underline{W}_c^2) : \{235.6,\ 2.829,\ 0.107,\ 0.014\} \tag*{[3.3.32a]}$$

and

$$\sigma_i(\underline{W}_o^2) : \{4.823,\ 0.023,\ 0.0002,\ 0\} \tag*{[3.3.32b]}$$

Therefore, the condition number of this state space model with respect to pointwise state control is 17,000. The condition number of the system with respect to zero input state observation is $3.6\text{x}10^8$. The controllability and observability matrices for this system are

$$\underline{P} = \begin{bmatrix} 1 & 3.2 & 6.7 & 10.6 \\ 0 & 1 & 3.3 & 6.7 \\ 0 & 0 & 1 & 3.2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \underline{Q} = \begin{bmatrix} 0.07 & 0.03 & -0.07 & 0.02 \\ 0.24 & -0.35 & 0.19 & -0.04 \\ 0.44 & -0.85 & 0.59 & -0.15 \\ 0.64 & -1.35 & 1.02 & -0.27 \end{bmatrix} \tag*{[3.3.33]}$$

The determinant of $\underline{P}$ and $\underline{Q}$ are

$$\det(\underline{P}) = 1.0 \tag*{[3.3.34a]}$$

$$\det(\underline{Q}) = 5.9\text{x}10^{-7} \tag*{[3.3.34b]}$$

The determinants of $\underline{P}$ and $\underline{Q}$ indicate that the system is controllable but almost unobservable. Examination of the singular values of $\underline{W}_c^2$ and $\underline{W}_o^2$

indicates that one state is unobservable and at least one state is almost uncontrollable. The large condition numbers for the controllability and observability Grammians indicate that this system would be difficult to control and to observe.

Example 3.9: If the state space system in example 3.7 is discretized with a sample time of 0.1, including a zero order hold, the resulting discrete model is

$$\underline{A} = \begin{bmatrix} 1.284 & 0.474 & -0.175 \\ -0.474 & 1.041 & -0.567 \\ -0.175 & 0.657 & 0.155 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} 0.411 \\ 0.282 \\ 0.061 \end{bmatrix} \quad \underline{C}^T = \begin{bmatrix} 0.411 \\ -0.282 \\ -0.061 \end{bmatrix} \tag{3.3.35}$$

the controllability and observability Grammians are given by

$$\underline{W}_c^2 = \begin{bmatrix} 1.456 & 0 & 0 \\ 0 & 0.148 & 0 \\ 0 & 0 & 0.011 \end{bmatrix} = \underline{W}_o^2 \tag{3.3.36}$$

and the controllability and observability matrices are given by

$$\underline{P} = \underline{Q}^T = \begin{bmatrix} -0.411 & 0.673 & 0.913 \\ 0.282 & 0.133 & -0.225 \\ -0.061 & 0.078 & -0.030 \end{bmatrix} \tag{3.3.37}$$

where

$$\{\sigma_i(\underline{W}_c^2) = \sigma_i\underline{W}_o^2) : 1.456, 0.148, 0.011\} \tag{3.3.38}$$

and

$$\det(\underline{P}) = \det(\underline{Q}) = 0.048 \tag{3.3.39}$$

Therefore

$$\gamma(\underline{W}_c) = \gamma(\underline{W}_o) = 135.2 \tag{3.3.40}$$

In this case, the system is as controllable as it is observable since the condition number of the controllability Grammian and the observability Grammian are equal. These realizations are known as balanced realizations which will be discussed in chapter 4. This system will be easier to control and to observe than the system in example 3.8 since the condition number of the Grammians is considerably smaller. Also, all the states should be reasonably controllable and observable.

3.4 Summary

Principal component analysis can be used to describe the shape of the subspace for the linear transformation corresponding to the convolution model between an input sequence and an output sequence for a dynamic process. As the sequence lengths increase, more process information is captured in the linear transformation causing the ellipsoid representing the system to expand. At some point, all the information will be captured in the system. Any further increase in the sequence length will not significantly expand the ellipsoid for the transformation.

The principal component vectors (singular vectors) and the component magnitudes (singular values) of the controllable and observable subspace provide a geometric description of the controllability and observability of the process. The principal components of these subspaces indicate explicitly whether a system is highly controllable/observable or weakly controllable/observable. Also, the relative magnitudes of the principal components describe the conditioning of the system with respect to controllability and observability and the sensitivity of the system to noise or modelling errors.

Chapter 4 Internally Balanced State Space Process Models

4.0 Introduction

There are several standard forms for state space models which are related through linear transformations. For each input/output model, there can be several state space models which will provide the same response to a given input sequence. Moore (1979a, 1981) introduced internally balanced state space realizations for continuous systems. In a balanced realization, the system states are as controllable as they are observable with the state vector ordered from the most controllable/observable state, $x_1(t)$, to the least controllable/observable state, $x_n(t)$. Laub (1980) and Silverman (1980) provided some insight into the stability of these models as well as the discrete time equivalent to Moore's initial theory. Several researchers have utilized internally balanced state space models in process model identification and model order reduction algorithms since they exhibit good numerical properties.

The following sections describe the process for obtaining balanced realizations from a minimal realization of the process. In addition, the properties of balanced realizations are discussed.

4.1 Continuous Time Models

Let $(\underline{A},\underline{B},\underline{C})$ represent a minimal realization of the continuous system given in equations 3.1.17a and 3.1.17b. If $\underline{T}$ is a matrix representing a linear transformation which scales the internal states such that:

$$\hat{\underline{x}}(t) = \underline{T}\underline{x}(t) \qquad [4.1.1]$$

the scaled system will be given by equations 3.1.26a,b,c,d. The controllability and observability Grammians of the "scaled" system are given by (Moore, 1979a, Laub, 1980):

$$\underline{\bar{W}}_c^2 = \underline{T}^{-1}\underline{W}_c^2(\underline{T}^{-1})^T \qquad [4.1.2a]$$

$$\underline{\bar{W}}_o^2 = \underline{T}^T\underline{W}_o^2\underline{T} \qquad [4.1.2b]$$

where:

$$\underline{W}_c^2 = \underline{Z}_c\underline{\Sigma}_c^2\underline{Z}_c^T = \left(\underline{Z}_c\underline{\Sigma}_c\right)\cdot\left(\underline{Z}_c\underline{\Sigma}_c\right)^T \qquad [4.1.3a]$$

$$\underline{W}_o^2 = \underline{Z}_o\underline{\Sigma}_o^2\underline{Z}_o^T = \left(\underline{Z}_o\underline{\Sigma}_o\right)\cdot\left(\underline{Z}_o\underline{\Sigma}_o\right)^T \qquad [4.1.3b]$$

are the controllability and observability Grammians of the original minimal realization $(\underline{A},\underline{B},\underline{C},\underline{D})$ in equations 3.1.17a and 3.1.17b. Let a matrix, $\underline{H}$, be defined as:

$$\underline{H} = \left[\underline{Z}_o\underline{\Sigma}_o\right]^T\left[\underline{Z}_c\underline{\Sigma}_c\right] = \underline{Z}_h\underline{\Sigma}_h\underline{V}_h^T \qquad [4.1.4]$$

and a linear transformation $\underline{T}_k$ be defined as:

$$\underline{T}_k = \underline{Z}_c\underline{\Sigma}_c\underline{V}_h\underline{\Sigma}_h^{-k} \qquad [4.1.5]$$

(Laub, 1980). If k=0, it can be shown that:

$$\underline{\bar{W}}_c^2 = \underline{T}_0^{-1}\ \underline{W}_c^2\underline{T}_0 = \underline{I} \qquad [4.1.6a]$$

and:

$$\underline{\bar{W}}_o^2 = \underline{T}_0^T\ \underline{W}_o^2\underline{T}_0 = \underline{\Sigma}_h^2 \qquad [4.1.6b]$$

Although all the states in this realization will be controllable, they may not be strongly observable. This type of realization is referred to as an input-normalized realization (Moore, 1979a, Laub 1980, Silverman 1980). If there are small components in $e^{\underline{A}^T t}\underline{C}^T$ which are offset by large components of $e^{\underline{A}t}\underline{B}$, a transformation to an input normalized realization will make all components of $e^{\underline{A}t}\underline{B}$ of unit magnitude and allow comparison of the components in $e^{\underline{A}^T t}\underline{C}^T$.

If k = 1 in equation 4.1.5, the controllability and observability Grammians for the resultant realization will be given by:

$$\underline{\hat{W}}_c^2 = \underline{T}_1^{-1} \underline{W}_c^2 \underline{T}_1 = \underline{\Sigma}_h^2 \qquad [4.1.7a]$$

and

$$\underline{\hat{W}}_o^2 = \underline{T}_1^T \underline{W}_o^2 \underline{T}_1 = \underline{I} \qquad [4.1.7b]$$

This type of realization is referred to as an output-normalized realization where all the states are observable. However, the state may not be strongly controllable (Moore, 1979a). If the components of $e^{\underline{A}t}\underline{B}$ are small in the minimal realization relative to $e^{\underline{A}^T t}\underline{C}^T$, they can be compared to each other in an output-normalized realization since all the components of $e^{\underline{A}^T t}\underline{C}^T$ will have a magnitude of one.

If k = 1/2, an internally balanced realization is obtained with the controllability and observability Grammians given by:

$$\underline{\hat{W}}_c^2 = \underline{T}_{1/2}^{-1} \underline{W}_c \underline{T}_{1/2} = \Sigma_h \qquad [4.1.8a]$$

and

$$\underline{\hat{W}}_o^2 = \underline{T}_{1/2}^T \underline{W}_o \underline{T}_{1/2} = \underline{\Sigma}_h \qquad [4.1.8b]$$

In this case, the magnitudes of the components of $e^{\underline{A}^T t}\underline{C}^T$ and $e^{\underline{A}t}\underline{B}$ will be equal since the Grammians are equal and diagonal. All the states will be as controllable as they are observable since the controllable and observable subspaces are equivalent. The component magnitudes of the controllable and observable subspaces ($\sigma_i(\underline{H})$; i=1,2,..,n) indicate the controllability/observability of the corresponding state.

Alternatively, the linear transformation $\underline{T}_k$ where:

$$\underline{T}_k = \underline{Z}_o \underline{\Sigma}_o^{-1} \underline{Z}_h \underline{\Sigma}_h^k \qquad [4.1.19]$$

can be used to obtain input-normalized (k=1), output-normalized (k=0) or internally balanced (k=1/2) realizations (Laub, 1980).[1] The linear transformations required to transform a system between these three realizations are aligned and differ only by scaling factors (Moore, 1979a):

$$\underline{T}_{ib} = \underline{T}_{on} \underline{\Sigma}_h^{1/2} = \underline{T}_{in} \underline{\Sigma}_h^{-1/2} \qquad [4.1.10]$$

Example 4.1: For the system in example 3.1 (equation 3.1.28), the matrix $\underline{H}$ in equation 4.1.4 is:

$$\underline{H} = \begin{bmatrix} -0.049 & 0.057 & -0.301 & -0.485 \\ -0.063 & -0.099 & 0.036 & -0.013 \\ 0.003 & -0.052 & -0.004 & 0.001 \\ 0.112 & -0.0002 & 0.0001 & -0.0001 \end{bmatrix} \qquad [4.1.11]$$

where

$$\underline{Z}_H = \begin{bmatrix} -1.000 & 0.008 & 0.027 & -0.006 \\ 0.021 & 0.777 & 0.476 & -0.411 \\ 0.008 & 0.186 & 0.451 & 0.873 \\ 0.017 & -0.601 & 0.755 & -0.262 \end{bmatrix} \qquad [4.1.12a]$$

$$\underline{V}_H = \begin{bmatrix} 0.087 & -0.790 & 0.606 & -0.044 \\ -0.104 & -0.583 & -0.764 & -0.258 \\ 0.524 & 0.167 & 0.082 & -0.832 \\ 0.841 & -0.094 & -0.208 & 0.490 \end{bmatrix} \qquad [4.1.12b]$$

$$\underline{\Sigma}_H = \begin{bmatrix} 0.576 & 0 & 0 & 0 \\ 0 & 0.148 & 0 & 0 \\ 0 & 0 & 0.090 & 0 \\ 0 & 0 & 0 & 0.019 \end{bmatrix} \qquad [4.1.12c]$$

If $\underline{T} = \underline{Z}_c \Sigma_c \underline{V}_h$, an input normalized realization is obtained where:

$$\underline{A}_{in} = \begin{bmatrix} -0.518 & 0.734 & 0.155 & -0.064 \\ -2.867 & -2.195 & -3.721 & 0.440 \\ 0.988 & 6.072 & -0.630 & 0.552 \\ 1.917 & 3.375 & -2.594 & -1.657 \end{bmatrix} \quad \underline{B}_{in} = \begin{bmatrix} -1.018 \\ -2.095 \\ 1.122 \\ 1.820 \end{bmatrix} \quad \underline{C}^T_{in} = \begin{bmatrix} -0.587 \\ 0.309 \\ 0.104 \\ -0.035 \end{bmatrix}$$

[4.1.13]

The controllability and observability Grammians are:

$$\underline{W}^2_{oin} = \begin{bmatrix} 0.332 & 0 & 0 & 0 \\ 0 & 0.022 & 0 & 0 \\ 0 & 0 & 0.008 & 0 \\ 0 & 0 & 0 & 0.0004 \end{bmatrix} \quad \underline{W}^2_{cin} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

[4.1.14]

Therefore, $\underline{W}^2_{oin} = \underline{\Sigma}^2_h$ and the realization $(\underline{A}_{in}, \underline{B}_{in}, \underline{C}_{in})$ is input normalized.

If $\underline{T} = \underline{Z}_c\underline{\Sigma}_c\underline{V}_h\underline{\Sigma}_h^{-1}$, an output normalized realization is obtained where:

$$\underline{A}_{on} = \begin{bmatrix} -0.518 & 2.867 & 0.988 & -1.917 \\ -0.734 & -2.195 & -6.072 & 3.375 \\ 0.155 & 3.721 & -0.630 & 2.594 \\ 0.064 & 0.440 & -0.552 & -1.657 \end{bmatrix} \quad \underline{B}_{on} = \begin{bmatrix} -0.587 \\ -0.309 \\ 0.101 \\ 0.035 \end{bmatrix} \quad \underline{C}^T_{on} = \begin{bmatrix} -1.018 \\ 2.095 \\ 1.122 \\ -1.820 \end{bmatrix}$$

[4.1.15]

The controllability and observability Grammians are:

$$\underline{W}^2_{con} = \begin{bmatrix} 0.332 & 0 & 0 & 0 \\ 0 & 0.022 & 0 & 0 \\ 0 & 0 & 0.008 & 0 \\ 0 & 0 & 0 & 0.0004 \end{bmatrix} \quad \underline{W}^2_{oon} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

[4.1.16]

Therefore, $\underline{W}^2_{oin} = \underline{\Sigma}^2_h$ and the realization $\{\underline{A}_{on}, \underline{B}_{on}, \underline{C}_{on}\}$ is output normalized.

If $\underline{T} = \underline{Z}_c\Sigma_c\underline{V}_h\Sigma_h^{-1/2}$, an internally balanced state space realization is obtained where:

$$\underline{A}_{ib} = \begin{bmatrix} -0.518 & 1.450 & 0.391 & -0.350 \\ -1.450 & -2.195 & -4.753 & 1.218 \\ 0.391 & 4.753 & -0.630 & 1.196 \\ 0.350 & 1.218 & -1.196 & -1.657 \end{bmatrix} \quad \underline{B}_{in} = \begin{bmatrix} -0.773 \\ -0.805 \\ 0.337 \\ 0.252 \end{bmatrix} \quad \underline{C}^T_{in} = \begin{bmatrix} -0.773 \\ 0.805 \\ 0.337 \\ -0.252 \end{bmatrix}$$

[4.1.17]

The controllability and observability Grammians are

$$\underline{W}^2_{cib} = \begin{bmatrix} 0.576 & 0 & 0 & 0 \\ 0 & 0.148 & 0 & 0 \\ 0 & 0 & 0.090 & 0 \\ 0 & 0 & 0 & 0.019 \end{bmatrix} = \underline{W}^2_{cib} \qquad [4.1.18]$$

Therefore, $\underline{W}^2_{cib} = \underline{W}^2_{oib} = \underline{\Sigma}(\underline{H})$ and the realizations $\{\underline{A}_{ib}, \underline{B}_{ib}, \underline{C}_{ib}\}$ is internally balanced. The above realizations were calculated using programs inpnorm.m, outnorm.m and intbal.m written in matlab which are given in Appendix C.

Although the matrices $\underline{H}$, $\underline{W}_c^2$ and $\underline{W}_o^2$ depend on the coordinate system of the state space, the singular values of $\underline{H}$ are invariant under linear transformation. $\underline{W}_o^2\underline{W}_c^2$ is also invariant under the linear transformations $\underline{T}_k$ since $\tilde{\underline{W}}_o^2\tilde{\underline{W}}_c^2 = \underline{\Sigma}_h^2$ for k equal to 0, 0.5 and 1 in equation 4.1.5 and 4.1.9. The singular values of $\underline{H}$:

$$\{\sigma_i(\underline{H}):\ i=1,2,\ldots,n\} \qquad [4.1.19]$$

referred to as second order modes of the system by Moore (1979a), are characteristic of the input and output properties of the system and not the specific state space structure. On the other hand, the singular vectors of $\underline{H}$, $\underline{W}_c^2$ and $\underline{W}_o^2$ are dependent on the coordinate system.

Example 4.2: For the internally balanced, input normalized and output normalized realizations given in example 4.1, the corresponding $\underline{H}$ matrices are given by:

$$\underline{H}_{in} = \begin{bmatrix} -0.135 & 0 & -0.271 & 0.137 \\ 0.019 & 0.003 & -0.106 & -0.0005 \\ 0.002 & -0.008 & 0 & 0.001 \\ 0.0001 & -0.008 & 0.0001 & 0.0003 \end{bmatrix} \qquad [4.1.20]$$

$$\underline{H}_{on} = \begin{bmatrix} 0.033 & -0.007 & 0.004 & 0.0003 \\ -0.086 & 0.015 & 0.005 & 0 \\ 0.319 & 0.004 & 0.001 & -0.0001 \\ 0.012 & 0.013 & -0.004 & 0.0002 \end{bmatrix} \qquad [4.1.21]$$

$$\underline{H}_{ib} = \begin{bmatrix} 0.332 & 0 & 0 & 0 \\ 0 & -0.022 & 0 & 0 \\ 0 & 0 & -0.008 & 0 \\ 0 & 0 & 0 & -0.0004 \end{bmatrix} \qquad [4.1.22]$$

The singular values for these matrices are:

$$\{\sigma_i(\underline{H}_{in}):\ 0.332,\ 0.022,\ 0.008,\ 0.0004\} \qquad [4.1.23a]$$

$$\{\sigma_i(\underline{H}_{on}):\ 0.332,\ 0.022,\ 0.008,\ 0.0004\} \qquad [4.1.23b]$$

$$\{\sigma_i(\underline{H}_{ib}):\ 0.332,\ 0.022,\ 0.008,\ 0.0004\} \qquad [4.1.23c]$$

Moore (1979a) and Pernebo (1982) showed that if the minimal realization, $(\underline{A},\underline{B},\underline{C})$, of the system is asymptotically stable, the internally balanced, input-normalized and output-normalized state space realizations will always be asymptotically stable:

$$e^{\underline{A}t} < 1 \qquad \text{for all } t>0 \qquad [4.1.24]$$

and

$$\underline{x}(t)\ ^{2} \rightarrow 0 \text{ as } t \rightarrow \infty \qquad [4.1:25]$$

In addition, if the minimal realization is asymptotically stable, the three realizations will be uniquely defined within an arbitrary unitary transformation. If the second order modes, $\sigma_i(\underline{H})$, are distinct and real, the basis vectors which define the internally balanced state space model will be unique within a change of sign (Moore, 1979a).

The condition number of the controllability and the observability Grammians can be used to assess the controllability and the observability of equivalent state space realizations. If the condition number of the Grammians is small, all the states can be said to be well controllable and observable. Any small modeling errors or perturbations in the states should not significantly affect the input-output properties of the state space model. If the condition numbers are large, the controllability and observability ellipsoids are highly distorted indicating an internal scaling imbalance where some of the states are almost uncontrollable and/or unobservable as indicated by the singular values of the Grammians. Appropriate scaling of the internal system via an appropriate linear transformation can change the shape of the ellipsoids and improve its condition number. For an input-normalized realization, the condition number of the controllability Grammian will be unity whereas the observability Grammian will depend on the relative magnitudes of the second order modes.

Similarly, the condition number of the observability Grammian is unity for the output-normalized realization and the condition number of the controllability Grammian depends on $\sigma_i(\underline{H})$. Moore (1979a) showed that if the state space model is an "internally balanced realization", the condition numbers with respect to controllability and zero input state observation will be the same and minimal. Therefore, a balanced realization provides optimal scaling of the system states with respect to controllability and observability.

Example 4.3: The condition numbers for the realizations given in example 3.1 and 4.1 were calculated within MATLAB. Table 4.1 summarizes the condition numbers with respect to controllability and observability for the various realizations of this system. The internally balanced model has a lower condition number with respect to controllability than an input normalized realization. The internally balanced model has a lower condition number with respect to observability than an output normalized realization.
The phase canonical realization is poorly conditioned with respect to both controllability and observability. Therefore, if the controllability and the observability of a system is important to analyze, an internally balanced realization will be the best conditioned system.

Shokoohi (1987) generalizes the concept of internally balanced realizations to "block balanced" realization. A state space realization (A,B,C,D) is said to be balanced in the wide sense if:

$$\underline{W}_c = \underline{W}_o = \underline{S} = \left[\begin{array}{c|c} \underline{S}_{11} & \underline{S}_{12} \\ \hline \underline{S}_{12} & \underline{S}_{22} \end{array}\right] \qquad [4.1.26]$$

Table 4.1: Condition Number with Respect to Controllability and Observability of the Realizations in Example 3.1 and 4.1

Realization	$\gamma(\underline{W}_c^2)$	$\gamma(\underline{W}_o^2)$
input normalized	30.0	1.0
output normalized	1.0	30.0
internally balanced	5.48	5.48
phase canoniacal	41.1	196.4

where $\underline{S}_{11} \in \mathbb{R}^{kxk}$ (Shokoohi, 1987). If $\underline{S}_{12} = \underline{0}$, the realization is said to be block balanced. If $\underline{S}_{12} = \underline{0}$ and $\underline{S}_{11}$ and $\underline{S}_{22}$ are diagonal, the realization $(\underline{A},\underline{B},\underline{C},\underline{D})$ is internally balanced. Therefore, an internally balanced realization is a subset of a block balanced realization. Shokoohi (1987) shows that all the properties of internally balanced realizations are preserved in block balanced realizations. In a similar manner, realizations which are input (output) normal in the wide sense and input (output) normal in the block partitioned sense are defined, respectively, as:

$$\underline{W}_c = \underline{I} \; ; \; \underline{W}_o = \underline{S} = \left[\begin{array}{c|c} \underline{S}_{11} & \underline{S}_{12} \\ \hline \underline{S}_{12} & \underline{S}_{22} \end{array}\right] \qquad [4.1.27]$$

$$\underline{W}_c = \underline{I}; \; \underline{W}_o = \underline{S} = \left[\begin{array}{c|c} \underline{S}_{11} & \underline{0} \\ \hline \underline{0} & \underline{S}_{22} \end{array}\right] \qquad [4.1.28]$$

Computation of block balanced realizations requires recursive solution of a nonsymmetric, algebraic Ricatti equation. Although the block balanced realizations are more general than the internally balanced realizations and possess the same properties, they can be more computationally intensive to obtain from minimal realizations than internally balanced realizations.

4.2 Discrete Time Models

The derivation of an internally balanced state space model from a minimal realization of a discrete time system is analogous to the continuous time case. Let $\underline{T}$ be a matrix representing a linear transformation of the discrete states, given in equations 3.1.19a and 3.1.19b, such that:

$$\bar{x}(k) = \underline{T}x(k) \qquad [4.2.1]$$

The transformed or scaled realization becomes:

$$\bar{\underline{x}}(k+1) = \underline{\bar{A}}\bar{\underline{x}}(k) + \underline{\bar{B}}\underline{u}(k) \qquad [4.2.2a]$$

$$\underline{y}(k) = \underline{\bar{C}}\bar{\underline{x}}(k) + \underline{D}\underline{u}(k) \qquad [4.2.2b]$$

where

$$\underline{\bar{A}} = \underline{T}^{-1}\underline{A}\underline{T} \qquad [4.2.2c]$$

$$\underline{\bar{B}} = \underline{T}^{-1}\underline{B} \qquad [4.2.2d]$$

$$\underline{\bar{C}} = \underline{C}\underline{T} \qquad [4.2.2e]$$

The discrete time controllability and observability Grammians for the minimal realization are given by equations 3.3.22a and 3.3.22b, respectively, and the scaled Grammians are given by:

$$\underline{\bar{W}}_c^2 = \underline{T}^{-1}\underline{W}_c^2(\underline{T}^{-1})^T \qquad [4.2.3a]$$

$$\underline{\bar{W}}_o^2 = \underline{T}^T\underline{W}_o^2\underline{T} \qquad [4.2.3b]$$

Input-normalized, output-normalized and internally balanced realizations are obtained from the matrix $\underline{H}$, given in equations 4.1.4, and the transformation matrix $\underline{T}_k$ in equation 4.1.5 (or 4.1.9) with k=0,1/2,1 respectively. The resulting input-normalized, output-normalized or internally balanced realization will be asymptotically stable ($\underline{A} \leq 1$) if and only if $\underline{\Sigma}_h$ has distinct diagonal elements and the minimal realization was asymptotically stable (Pernebo, 1982). This condition represents a weaker result than the stability condition for the corresponding continuous system. It should be noted that $\underline{A} = 1$ only if the realization is stable.

Moore (1979a, 1981) showed that the matrix $\underline{H}$ for a discretized system, given in equation 4.1.4, can be obtained from a process response by sampling arbitrarily fast over the time interval $[0,\infty)$. However, $\underline{H}$ contains essentially the same information as the discrete impulse hankel matrix for the process, $\underline{\Gamma}$, since:

$$\sigma_i(\underline{H}) = \lim_{t_s \to 0} \sigma_i(\underline{\Gamma}) \qquad [4.2.4]$$

where t_s is the sampling time. A hankel matrix is a matrix which is constant along the antidiagonals. The impulse hankel matrix is composed of the impulse response parameters sampled at discrete intervals. If there are an infinite number of impulse response parameters, h_i, the (SISO) hankel matrix will be infinite and symmetric and given by:

$$\underline{\Gamma} = \begin{bmatrix} h_1 & h_2 & h_3 \,..h_i & . & . & .. \\ h_2 & h_3 & h_4 \,..h_{i+1} & . & . & .. \\ \vdots & \vdots & \vdots & & & \\ h_j & h_{j+1} & \ldots h_{i+j-1} & & . & .. \\ \vdots & \vdots & \vdots & & & \end{bmatrix} \qquad [4.2.5]$$

If the impulse response of a SISO process settles after 2N-1 impulse response parameters such that $h_i = 0$ for $t > (2N-1)^* t_s$, the corresponding impulse hankel matrix is given by:

$$\underline{\Gamma} = \begin{bmatrix} h_1 & h_2 & h_3 \cdots & h_{2N-1} \\ h_2 & h_3 & h_4 \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ h_{2N-2} & h_{2N-1} & \cdots\cdots & 0 \\ h_{2N-1} & 0 & \cdots\cdots & 0 \\ 0 & 0 & \cdots\cdots & 0 \\ \vdots & \vdots & & \vdots \end{bmatrix} \qquad [4.2.6]$$

An NXN impulse hankel matrix can be defined as:

$$\underline{\Gamma}_{N,N} = \begin{bmatrix} h_1 & h_2 & h_3 \cdots & h_N \\ h_2 & h_3 & h_4 \cdots & h_{N+1} \\ \vdots & \vdots & \vdots & \vdots \\ h_{N-1} & h_N & \cdots\cdots & h_{2N-2} \\ h_N & h_{N+1} & \cdots\cdots & h_{2N-1} \end{bmatrix} = \begin{bmatrix} \underline{c}\underline{b} & \underline{c}\underline{A}\underline{b} & \cdots\cdots & \underline{c}\underline{A}^{N-1}\underline{b} \\ \underline{c}\underline{A}\underline{b} & \underline{c}\underline{A}^{2}\underline{b} & \cdots\cdots & \underline{c}\underline{A}^{N}\underline{b} \\ \vdots & \vdots & & \vdots \\ & & \cdots\cdots & \\ \underline{c}\underline{A}^{N-1}\underline{b} & & \cdots\cdots & \underline{c}\underline{A}^{2N-2}\underline{b} \end{bmatrix}$$

$$\underline{\Gamma}_{N,N} = \begin{bmatrix} \underline{c} \\ \underline{c}\underline{A} \\ \vdots \\ \underline{c}\underline{A}^{N-1} \end{bmatrix} \cdot \begin{bmatrix} \underline{b} & \underline{A}\underline{b} & \cdots & \underline{A}^{N-1}\underline{b} \end{bmatrix} = \underline{Q}\cdot\underline{P} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T \qquad [4.2.7]$$

which contains all the information of the infinite impulse hankel matrix in equation 4.2.6.

The impulse hankel matrix can be used to predict the response of the system to a given past input sequence assuming no further control action is taken. Given an input sequence:

$$\{\ldots 0,0,\ldots u(t-N),\ldots,u(t-2),u(t-1),0,0,\ldots\} \qquad [4.2.8]$$

the future process output is given by

$$\begin{bmatrix} y(k) \\ y(k+1) \\ \vdots \\ y(k+N-1) \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & \cdots & h_N \\ h_2 & h_3 & \cdots & h_{N+1} \\ \vdots & \vdots & & \vdots \\ h_N & h_{N+1} & \cdots & h_{2N-1} \end{bmatrix} \cdot \begin{bmatrix} u(k-1) \\ u(k-2) \\ \vdots \\ u(k-N) \end{bmatrix}$$

$$\underline{y}(k) = \underline{\Gamma} \cdot \underline{u}(k) \qquad [4.2.9]$$

For a multivariable system with m outputs and p inputs, the generalized impulse hankel matrix for the process will be an infinite block matrix where $\underline{h}_i$ are m×p matrices:

$$\underline{\Gamma} = \begin{bmatrix} \underline{h}_1 & \underline{h}_2 & \underline{h}_3 & \cdots & \underline{h}_i & \cdots \\ \underline{h}_2 & \underline{h}_3 & \underline{h}_4 & \cdots & \underline{h}_{i+1} & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \\ \underline{h}_{j-1} & \underline{h}_j & \cdots & & \underline{h}_{i+j} & \cdots \\ h & \underline{h}_{j+1} & \cdots & & \underline{h}_{i+j-1} & \cdots \\ \vdots & \vdots & & & \vdots & \end{bmatrix} \qquad [4.2.10]$$

The block impulse hankel matrix will be symmetric if and only if all the $\underline{h}_i$ are symmetric mxp matrices.

An impulse hankel operator, $\underline{\Gamma}(t-\tau)$, can be defined for continuous systems which is analogous to the impulse hankel matrix for discrete time systems (Wahlberg, 1986, Glover, 1984, Bettayeb, 1980):

$$(\underline{\Gamma}u)(t) = \int_{-\infty}^{0} \underline{C}e^{\underline{A}(t-\tau)}\underline{B}\underline{u}(\tau)d\tau = \int_{0}^{\infty} \underline{C}e^{\underline{A}(t+\lambda)}\underline{B}u(-\lambda)d\lambda = \int_{0}^{\infty} \underline{h}(t+\lambda)u(-\lambda)d\lambda \qquad [4.2.11]$$

where:

$$y(t) = \int_{0}^{t} \underline{\Gamma}(t-\tau)u(\tau)d\tau \qquad [4.2.12]$$

Example 4.4: For the discrete system given in example 3.1 (equation 3.1.32) with a sampling time of 0.1:

$$\underline{H} = \begin{bmatrix} 1.910 & -0.944 & 0.456 & -0.189 \\ -1.136 & 0.185 & -0.008 & -0.008 \\ 0.176 & 0.003 & -0.003 & -0.0008 \\ -0.0008 & -0.0001 & 0 & 0 \end{bmatrix} \qquad [4.2.13]$$

with singular values:

$$\{\sigma_i(\underline{H}): 2.448,\ 0.380,\ 0.017,\ 0\} \qquad [4.2.14]$$

The resulting input-normalized, output normalized and internally balanced realizations, obtained from the Matlab programs, are:

$$\underline{A}_{in} = \begin{bmatrix} 1.122 & 0.136 & 0.008 & 0 \\ -0.873 & 0.901 & 0.075 & 0.0003 \\ 1.105 & -1.696 & 0.620 & 0.006 \\ -3.492 & 3.137 & -3.145 & 0.677 \end{bmatrix} \quad \underline{B}_{in} = \begin{bmatrix} 0.316 \\ 0.699 \\ 0.601 \\ 0.224 \end{bmatrix} \quad \underline{C}^T_{in} = \begin{bmatrix} 0.775 \\ -0.266 \\ 0.010 \\ 0 \end{bmatrix}$$

[4.2.15]

$$\underline{A}_{on} = \begin{bmatrix} 1.122 & 0.873 & 1.105 & 3.492 \\ -0.136 & 0.901 & 1.696 & 3.137 \\ 0.008 & -0.075 & 0.620 & 3.144 \\ 0 & 0.0003 & -0.006 & 0.677 \end{bmatrix} \quad \underline{B}_{on} = \begin{bmatrix} 0.775 \\ 0.266 \\ 0.010 \\ 0 \end{bmatrix} \quad \underline{C}^T_{on} = \begin{bmatrix} 0.316 \\ -0.699 \\ 0.601 \\ -0.224 \end{bmatrix}$$

[4.2.16]

$$\underline{A}_{ib} = \begin{bmatrix} 1.122 & 0.344 & 0.092 & 0.013 \\ -0.344 & 0.901 & 0.357 & 0.030 \\ 0.092 & -0.357 & 0.620 & 0.141 \\ -0.013 & 0.030 & -0.141 & 0.677 \end{bmatrix} \quad \underline{B}_{ib} = \begin{bmatrix} 0.495 \\ 0.431 \\ 0.078 \\ 0.001 \end{bmatrix} \quad \underline{C}^T_{ib} = \begin{bmatrix} 0.495 \\ -0.430 \\ 0.078 \\ -0.001 \end{bmatrix}$$

[4.2.17]

The impulse response for this system is shown in Figure 3.3. It appears that the settling time of the process to an impulse is approximately 50 time steps. The impulse hankel matrix for this system would be:

$$\underline{\Gamma} = \begin{bmatrix} 0.065 & 0.242 & 0.449 & 0.638 & \cdots \\ 0.242 & 0.449 & 0.638 & 0.770 & \cdots \\ 0.449 & 0.638 & 0.770 & 0.825 & \cdots \\ 0.638 & 0.770 & 0.825 & 0.779 & \cdots \\ 0.770 & 0.825 & 0.779 & 0.704 & \cdots \\ 0.825 & 0.779 & 0.704 & 0.568 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix} \qquad [4.2.18]$$

The singular values of the 45x45 impulse hankel matrix are:

$$\{\sigma_i(\Gamma): 6.544,\ 1.577,\ 0.942,\ 0.185,\ 0,\ 0,\ldots,0\} \qquad [4.2.19]$$

which are not exactly equal to the second order modes of the system. As the sampling time for the system decreases, the singular values of the impulse hankel matrix will approach the second order modes of the system.

Example 4.5: For the discrete system described by the transfer function model:

$$y(k) = \frac{0.053z^{-1} - 0.032z^{-2}}{1 - 1.684z^{-1} + 0.705z^{-2}} \cdot u(k) \qquad [4.2.20]$$

the corresponding state space model in controller canonical form (from Matlab) is:

$$\underline{A} = \begin{bmatrix} 1.684 & -0.705 \\ 1 & 0 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \underline{C} = \begin{bmatrix} 0.053 & -0.032 \end{bmatrix} \qquad [4.2.21]$$

The corresponding input-normalized, output normalized and internally balanced realizations are given by:

$$\underline{A}_{in} = \begin{bmatrix} 1.033 & 0.019 \\ -1.686 & 0.651 \end{bmatrix} \quad \underline{B}_{in} = \begin{bmatrix} 0.689 \\ 0.725 \end{bmatrix} \quad \underline{C}_{in} = \begin{bmatrix} 0.078 & -0.0009 \end{bmatrix} \qquad [4.2.22]$$

$$\underline{A}_{on} = \begin{bmatrix} 1.033 & 1.686 \\ -0.019 & 0.651 \end{bmatrix} \quad \underline{B}_{on} = \begin{bmatrix} 0.078 \\ 0.009 \end{bmatrix} \quad \underline{C}_{on} = \begin{bmatrix} 0.689 & -0.725 \end{bmatrix} \qquad [4.2.23]$$

$$\underline{A}_{ib} = \begin{bmatrix} 1.033 & 0.180 \\ -0.180 & 0.651 \end{bmatrix} \quad \underline{B}_{ib} = \begin{bmatrix} 0.232 \\ 0.026 \end{bmatrix} \quad \underline{C}_{ib} = \begin{bmatrix} 0.232 & -0.026 \end{bmatrix} \qquad [4.2.24]$$

The matrix $\underline{H}$ required to convert the controller canonical form to the equivalent input-normalized, output normalized and internally balanced realizations is given by:

$$\underline{H} = \begin{bmatrix} 0.107 & -0.036 \\ -0.003 & -0.0004 \end{bmatrix} \qquad [4.2.25]$$

with singular values

$$\{\sigma_i(\underline{H}): 0.113,\ 0.001\} \qquad [4.2.26]$$

Therefore, the second order modes of the system are 0.336 and 0.032. From the impulse response, shown in Figure 4.1, it appears that the process settles in t = 60. The singular values of the 60x60 impulse hankel matrix formed from the impulse response parameters are:

$$\{\sigma_i(\Gamma): 0.553,\ 0.043,\ 0.003,\ 0.002,\ 0.001,\ \ldots,\ 0\} \qquad [4.2.27]$$

Therefore, the system is approximately second order since $\sigma_2 \gg \sigma_3$.

Maciejowski (1985) defined k-balanced and ∞-balanced state space realizations for multivariable discrete time systems. A corresponding result for continuous time systems is not presented. A system is said to be k-balanced if $\underline{P}_k\underline{P}_k^T = \underline{Q}_k^T\underline{Q}_k = \underline{\Sigma}_h^2 = \underline{W}_c^2 = \underline{W}_o^2$ where

$$\underline{P}_k = \left[\underline{B}\quad \underline{AB}\quad \underline{A}^2\underline{B}\ \ldots\ldots\ \underline{A}^k\underline{B}\right] \qquad [4.2.28a]$$

$$\underline{Q}_k^T = \left[\underline{C}\quad \underline{CA}\quad \underline{CA}^2\ \ldots\ldots\ \underline{CA}^k\right]^T \qquad [4.2.28b]$$

are the controllability and observability matrices, respectively. A system is said to be ∞-balanced if $\underline{\Sigma}_h^2$ satisfied the discrete time Lyapunov equations:

$$\underline{\Sigma}_h^2 - \underline{A}\underline{\Sigma}_h^2\underline{A}^T = \underline{BB}^T \qquad [4.2.29a]$$

$$\underline{\Sigma}_h^2 - \underline{A}^T\underline{\Sigma}_h^2\underline{A} = \underline{C}^T\underline{C} \qquad [4.2.29b]$$

Maciejowski (1985) has differentiated between an internally balanced realization obtained from a finite impulse response and one obtained from an infinite impulse hankel matrix. Since $\underline{\Gamma} = \underline{Q}\,\underline{P}$, the impulse hankel matrix of the system given by $\underline{Q}_k$ and $\underline{P}_k$ will be kxk whereas the realization obtained from the Lyapunov equations will be infinite if the impulse hankel matrix is infinite.

4.3 Summary

From the examples presented in this chapter, it becomes apparent that a balanced realization provides optimal scaling of the state space with respect to controllability and observability. The condition number is equal and minimal for both the controllable subspace and the observable subspace. Since, the sensitivity of the balanced system to noise and modelling errors has also been minimized, state space controllers which are designed using balanced realizations should be relatively robust. However, if the main objective of the control scheme is to maximize observability, state space controllers designed using an input-normalized realization would be optimal. The condition number of the observable subspace for an input normalized realization is minimal whereas the condition number of the controllable subspace can be very large. Conversely, if the main objective of the control is to maximize controllability, state space controllers designed using an output-normalized realization would be optimal. In an output normalized realization, the condition number of the controllable subspace is minimal but the condition number of the observable subspace can be very large.

Chapter 5 Model Identification and Model Order Reduction Techniques Using Singular Value Analysis

5.0 Introduction

Control strategies which utilize model based controllers such as dynamic matrix control (DMC), generalized predictive control (GPC) or adaptive control schemes require an accurate process model to ensure robustness and stability and to achieve "good" performance. Accurate process models are also required to analyze and design multivariable control systems which will be discussed in chapter 6. The most common identification methods for advanced control implementations are one of the many variants of the recursive least squares methods (RLS) and time series analysis techniques. Shah (1987) presents a comprehensive review of recursive least squares routines. Anderson (1967) reviews time series analysis techniques. The least square routines require an apriori knowledge of the structure of the process model. However, in most practical applications, the actual order of the process is comparatively large and the structure is generally unknown. Time series analysis techniques can be used to identify a model structure prior to identifying the model parameters.

Recently, several researchers (Kung, 1978, Shokoohi, 1987a, Wahlberg, 1986, Silverman, 1980, Moore, 1979a, 1981, Gerstle, 1984, Al-Saggaf, 1987, Maciejowski, 1985 etc.) have utilized singular value analysis techniques to identify state space models from an impulse or step response of the process. In these methods, the model order is determined during the identification. Niederlinski (1984) and Verhaegen (1985) have used SVD to identify process models using a least squares identification routine (Niederlinski, 1984, Verhaegen, 1985).

Most parametric identification techniques are implemented in three steps:

i) choose a system order n

ii) identify a model from the input-output behaviour

iii) verify that the assumed n is reasonable

This type of approach to model identification can result in an iteration cycle since the system order, n, is generally unknown and the initial estimate of n may be inappropriate. Also, if the assumed model order or model structure is different from the true system order and structure, there will be a bias in the predicted process output relative to the actual output due to the modeling errors. Over parameterization increases the order of the model and increases the flexibility of the identified model to capture all the dynamic characteristics of the process. However, the matrices used in the identification can be large and nearly singular. Maciejowski (1985) and Wahlberg (1986) modified this conventional approach to eliminate the iteration step such that the identification algorithm requires only one pass through the process data. There are three steps in their approach:

i) choose the largest possible system order, N. The only limitations on the assumed order are CPU processing time, CPU load and memory size.

ii) identify a balanced realization of the system from the input-output behaviour using hankel matrix techniques or least squares techniques

iii) reduce the order of the balanced realization to a reasonable value

Although this approach does not require iteration using the input-output data of the process, the selection of an order for the final balanced realization may require some iteration.

Model order reduction techniques are used to generate low order process models which will approximate the actual structure and order of the process. The dominant frequency characteristics of the process should be retained in the reduced order model. The utilization of these reduced order models in the control algorithms will simplify any mathematical manipulations and therefore reduce the load on the computer. However, the model error between the actual process response and the model response to a given input or disturbance should be minimized to obtain robust control using the reduced order models. Several model reduction techniques have been developed which employ singular value analysis. Moore (1979a, 1981) introduced a method of model reduction using internally balanced state space models. Kung (1978), Glover (1984) and several other researchers have developed hankel norm model reduction techniques to obtain approximate and optimal state space realizations.

Several techniques for model identification and model order reduction which utilize singular value decomposition are described in this chapter. Through simulated examples, these methods are evaluated and recommendations for their use are given.

5.1 Hankel Matrix Model Identification Techniques

Several researchers have utilized the impulse hankel matrix and it's shift-invariance property to obtain minimal state space realizations and balanced realizations of the process. From the realization, a transfer function model, an impulse response model or a step response model can be generated. The general state space model identification procedure involves finding a linear transformation, $\underline{T}$, such that a set of controllable and observable state space models can be obtained from input-output behaviours (or transfer functions) (Maciejowski, 1985). Minimization of an error function between a full order model and a reduced order model is used to determine if the reduced model adequately approximates the original system behaviour. This type of identification algorithm does not require any apriori structural knowledge of the process. However, noise in the process can be a problem since the impulse response (or step response) is used for the identification. In the deterministic case, it is possible to find an exact minimal realization using the impulse hankel matrix. However, in the presence of noise, all the techniques will fail to produce a finite dimensional realization but an approximation can be obtained by utilizing SVD of the impulse hankel matrix. Kung (1978) and Damen (1982) present summaries of various impulse hankel matrix identification methods.

The N^{th} order partial realization problem is the determination of a discrete realization $\{\underline{A},\underline{B},\underline{C},\underline{D}\}$ from partial process information in the form of a finite discretized impulse (or step) response of the process such that

$$\underline{h}_i = \underline{C}\underline{A}^{i-1}\underline{B} \qquad i = 1,2,\ldots,N \qquad [5.1.1]$$

where $\{\underline{h}_i : i=1,2,\ldots,N\}$ is a finite sequence of impulse response parameters (Tether, 1970). $\{\underline{A},\underline{B},\underline{C},\underline{D}\}$ will be a minimal partial realization of order N if and only if the size of $\underline{A}$ is minimal among all possible partial realizations (Tether, 1970). A partial realization of order N is equivalent to matching the first N terms of the impulse response (Tether, 1970). An approximate realization is a realization which approximates a partial realization. Algorithms have been developed using the impulse hankel matrix of the system for time invariant processes (Silverman, 1980, Maciejowski, 1985, Kung, 1978) and time variant process (Shokoohi, 1983, 1984, 1987a). The following discussion will concentrate on discrete time systems. However, continuous time systems can be dealt with in the same manner since the identification problem is the same.

From Kronecker's theorem, the infinite sequence of impulse parameters $\{\underline{h}_i: i=1,2,...\}$ defines a finite minimal state space realization of order n if and only if the infinite impulse hankel matrix formed by the sequence has rank less than or equal to n (Kung, 1978, Kailath, 1980). The hankel matrix will have a finite rank n if and only if $A(z^{-1})$ has n stable poles (analytic outside the unit circle $|z|\geq 1$) where (Silverman, 1980):

$$H(z^{-1}) = B(z^{-1})/A(z^{-1}) \qquad [5.1.2]$$

If the system being identified is noisy, the calculated rank of $\underline{\Gamma}$ may not represent the actual order of the system as discussed previously in chapter 2. However, if the noise can be estimated, a noise criteria can be used to determine the approximate order of the process.

An infinite impulse hankel matrix, given in equation 4.1.33, can be formed from the infinite impulse response sequence $\{h_i: i=1,2,...\}$. If

$$\tau_s = N \times T_s \qquad [5.1.3]$$

where τ_s is the settling time of the process and T_s is the sampling time, the infinite impulse response can be truncated to $\{h_i: i=1,2,...,N\}$. Therefore, the rank of the finite impulse hankel matrix formed from the truncated sequence will approximate the order of the process. The truncation of the impulse response sequence at τ_s will introduce a error into the identification routine. However, if a sufficiently long impulse response sequence is obtained such that h_i is small for i>N, the error should be insignificant. The resulting impulse hankel matrix will be given by

$$\underline{\Gamma}_{N,N} = \begin{bmatrix} \underline{h}_1 & \underline{h}_2 & \underline{h}_3 & .. & \underline{h}_N \\ \underline{h}_2 & \underline{h}_3 & \underline{h}_4 & .. & \underline{h}_{N+1} \\ \vdots & \vdots & \vdots & & \vdots \\ \underline{h}_{N-1} & \underline{h}_N & & \ldots\ldots & \underline{h}_{2N-2} \\ \underline{h}_N & \underline{h}_{N+1} & & \ldots\ldots & \underline{h}_{2N-1} \end{bmatrix} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T = \underline{Q}_N\cdot\underline{P}_N \qquad [5.1.4a]$$

$$\underline{\Gamma}_{N,N} = \begin{bmatrix} \underline{C}\underline{B} & \underline{C}\underline{A}\underline{B} & \underline{C}\underline{A}^2\underline{B} \ldots & \underline{C}\underline{A}^{N-1}\underline{B} \\ \underline{C}\underline{A}\underline{B} & \underline{C}\underline{A}^2\underline{B} & \underline{C}\underline{A}^3\underline{B} \ldots & \underline{C}\underline{A}^N\underline{B} \\ \vdots & \vdots & \vdots & \vdots \\ \underline{C}\underline{A}^{N-2}\underline{B} & \ldots\ldots & \ldots\ldots & \underline{C}\underline{A}^{2N-2}\underline{B} \\ \underline{C}\underline{A}^{N-1}\underline{B} & & \ldots\ldots & \underline{C}\underline{A}^{2N-1}\underline{B} \end{bmatrix} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T = \underline{Q}_N\cdot\underline{P}_N \qquad [5.1.4b]$$

If $h_i \simeq 0$ for i>N, the impulse hankel matrix will be an upper triangular matrix. In this case, the size of the impulse hankel matrix can be reduced without losing any process information. If M=N/2, a full impulse hankel matrix can be formed as (Maciejowski, 1984):

$$\underline{\Gamma}_{M,M} = \begin{bmatrix} h_1 & h_2 & h_3 & .. & h_M \\ h_2 & h_3 & h_4 & .. & h_{M+1} \\ \vdots & \vdots & \vdots & & \vdots \\ h_{M-1} & h_M & & \ldots\ldots & h_{2M-2} \\ h_M & h_{M+1} & & \ldots\ldots & h_{2M-1} \end{bmatrix} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T \qquad [5.1.5]$$

The computation load required for the identification will be reduced by reducing the size of the impulse hankel matrix.

5.1.1 Ho-Kalman Algorithm and Variations

Ho and Kalman (1966) developed an algorithm to obtain minimal realizations from the discretized infinite impulse response for a finite dimensional, deterministic linear system. Since input/output data is used in the identification, an apriori knowledge of the system order (the number of state variables) is not required. If the dimension of the system is finite, the identification method will give an exact realization with the right dimension (Ho, 1966). Given the impulse hankel matrix, $\underline{\Gamma}_{N,N}$, in equation 5.1.4 formed from the impulse response of the system, a matrix $\underline{\Gamma}^{\uparrow}_{N,N}$ can be defined as

$$\underline{\Gamma}^{\uparrow}_{N,N} = \begin{bmatrix} h_2 & h_3 & h_4 & \cdots & h_{N+1} \\ h_3 & h_4 & h_5 & \cdots & h_{N+2} \\ \vdots & \vdots & \vdots & & \vdots \\ h_N & h_{N+1} & & \cdots\cdots & h_{2N-1} \\ h_{N+1} & h_{N+2} & & \cdots\cdots & h_{2N} \end{bmatrix} = \underline{Q}\cdot\underline{A}\cdot\underline{P} \qquad [5.1.6]$$

which is the original impulse hankel matrix shifted up by one (block) row with the appropriate impulse parameters entered into the bottom row. The shifted matrix utilizes one additional impulse response (block) parameter and retains the hankel structure. Two non-singular matrices, $\underline{R}$ and $\underline{S}$, are found such that

$$\underline{R}\cdot\underline{\Gamma}_{N,N}\cdot\underline{S} = \begin{bmatrix} \underline{I}_n & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} = \underline{U}_n^T\cdot\underline{U}_n = \underline{J} \qquad [5.1.7]$$

where $\underline{I}_n$ is an nxn identity matrix, n is the rank of $\Gamma_{N,N}$ ($n \leq N$) and

$$\underline{U}_n = [\ \underline{I}_n \quad \underline{0}\ ] \qquad [5.1.8]$$

The resulting n^{th} order minimal realization is given by

$$\underline{A} = \underline{U}_n\cdot[\underline{J}\cdot\underline{R}\cdot\underline{\Gamma}^{\uparrow}_{N,N}\cdot\underline{S}\cdot\underline{J}\ ]\cdot\underline{U}_n^T \qquad [5.1.9a]$$

$$\underline{B} = \underline{U}_n\cdot[\underline{J}\cdot\underline{R}\cdot\underline{\Gamma}_{N,N}\cdot\underline{E}_p^T] \qquad [5.1.9b]$$

$$\underline{C} = [\underline{E}_m\cdot\underline{\Gamma}_{N,N}\cdot\underline{S}\cdot\underline{J}\ ]\cdot\underline{U}_n^T \qquad [5.1.9c]$$

where

$$\underline{E}_p = [\ \underline{I}_p,\ \underline{0}_p] \quad \in \mathbb{R}^{pxN} \qquad [5.1.9d]$$

$$\underline{E}_m = [\ \underline{I}_m, \underline{0}_m] \quad \in \mathbb{R}^{m\mathrm{x}N} \qquad [5.1.9e]$$

are block matrices for the system with p inputs and m outputs. $\underline{E}_p^T$ and $\underline{E}_m$ selects the first p columns (number of inputs) and the first m rows (number of outputs) of $\underline{\Gamma}_{N,N'}$, respectively. The Ho-Kalman algorithm will identify a unique, finite realization of dimension n from the infinite impulse response if and only if there exists integers s and s' such that (Ho, 1966, Tether, 1970):

$$\mathrm{rank}(\underline{\Gamma}_{s,s'}) = \mathrm{rank}(\underline{\Gamma}_{s+i,s'+j}) \qquad \begin{array}{l} \forall\ i = 0,1,2,... \\ \forall\ j = 0,1,2,... \end{array} \qquad [5.1.10]$$

where s'≤ n and s ≤ n. Ho (1966) shows that this realization is minimal and every minimal realization can be obtained by suitably choosing $\underline{R}$ and $\underline{S}$.

Tether (1970) showed that the Ho-Kalman algorithm could be used to approximate an infinite dimensional deterministic system by a finite dimensional minimal partial realization. A partial realization will always exist via the Ho-Kalman algorithm since every finite sequence of impulse response parameters $\{h_i : i=1,2,...,N\}$ admits an extension sequence $\{h_i : i=N+1,N+2,...\}$ for which a completely controllable and observable partial realization exists of order N (Lemma 1, Tether, 1970). A minimal partial realization will always exist since the set of dimensions less than N is finite and the minimum can always be attained (Tether, 1970). A unique minimal partial realization will exist if and only if the extension sequence is unique (Tether 1970). The dimension of the minimal partial realization is given by (Theorem 1, Tether 1970):

$$n = \sum_{i=1}^{N} \mathrm{rank}\ (\ \underline{\Gamma}_{i,N+1-i}\ ) - \sum_{i=1}^{N} \mathrm{rank}\ (\ \underline{\Gamma}_{i,N-i}\ ) \qquad [5.1.11]$$

The resulting realization will only approximate the transient response of the system.

Zeigler (1974) used a modified Ho-Kalman algorithm to obtain approximate realizations of a given dimension for stochastic systems subject to a constraint on the state space dimension. In the presence of noise, the calculated rank of the impulse hankel matrix may not reflect the true dimension of the process. Therefore, to employ the Ho-Kalman identification algorithm to a noisy system, the impulse hankel matrix in equation 5.1.4 formed from the input/output data must be factored through a space of dimension r such that the resulting matrix of rank r approximates the impulse hankel matrix of the system impulse response. Zeigler (1974) used SVD of the impulse hankel matrix. In this case, the effect of noise on the calculation of matrix rank can be "filtered out" by choosing r based on the magnitude of the singular values of $\underline{\Gamma}_{N,N}$. The accuracy of the identified model will depend on the process noise and the choice of r.

If $\underline{\Gamma}_{N,N}$ is the impulse hankel matrix formed from the finite set of impulse response parameters $\{h_i : i=1,2,..,N\}$ where n is the rank $\underline{\Gamma}_{N,N}$, the best r^{th} rank approximation, in the least squares sense, to $\underline{\Gamma}_{N,N}$ is a matrix, $\underline{\Gamma}_{r,r}$, formed from $\underline{\Gamma}_{N,N}$ with the smallest (N-r) singular values set equal to zero:

$$\underline{\Gamma}_{N,N} = \begin{bmatrix} \underline{Z}_r & \underline{Z}_2 \end{bmatrix} \cdot \begin{bmatrix} \underline{\Sigma}_r & 0 \\ 0 & \underline{\Sigma}_2 \end{bmatrix} \cdot \begin{bmatrix} \underline{V}_r^T \\ \underline{V}_2^T \end{bmatrix} = \underline{Z} \cdot \underline{\Sigma}^{1/2} \cdot \underline{\Sigma}^{1/2} \cdot \underline{V}^T = \underline{Q} \cdot \underline{P} \qquad [5.1.12]$$

$$\underline{\Gamma}_{r,r} = \begin{bmatrix} \underline{Z}_r & \underline{Z}_2 \end{bmatrix} \cdot \begin{bmatrix} \underline{\Sigma}_r & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \cdot \begin{bmatrix} \underline{V}_r^T \\ \underline{V}_2^T \end{bmatrix} = \underline{Z}_r \cdot \underline{\Sigma}_r \cdot \underline{V}_r^T = \underline{Q}_r \cdot \underline{P}_r^T \qquad [5.1.13]$$

The set of singular values $(\sigma_i : i = 1,2,...,N)$ indicates the error between the full matrix, $\underline{\Gamma}_{N,N}$, and the matrix of lower rank, $\underline{\Gamma}_{r,r}$, in terms of the spectral norm (Kung, 1978, Glover, 1984, Silverman, 1980)

$$\underline{\Gamma}_{N,N} - \underline{\Gamma}_{r,r} = \sigma_{r+1}(\underline{\Gamma}_{N,N}) \qquad [5.1.14]$$

If $\sigma_r(\underline{\Gamma}_{N,N}) \gg \sigma_{r+1}(\underline{\Gamma}_{N,N})$, the approximate order of the system is 'r' although the calculated rank of $\underline{\Gamma}_{N,N}$ is 'n'. From equation 5.1.6,

$$\underline{A} = \underline{Q}^{-1} \cdot \underline{\Gamma}^{\uparrow}_{N,N} \cdot \underline{P}^{-1} \qquad [5.1.15]$$

Applying Ho-Kalman's algorithm, the best rth order realization of a system with p inputs and m outputs, in the least squares sense, is given by

$$\underline{B}_r = 1^{st} \text{ r rows X } 1^{st} \text{ p cols of } \underline{P}_r^T = \underline{\Sigma}_r^{-1/2}\underline{Z}_r^T\underline{\Gamma}_{N,N}\underline{E}_p^T \qquad [5.1.16a]$$

$$\underline{C}_r = 1^{st} \text{ m rows X } 1^{st} \text{ r cols of } \underline{Q}_r = \underline{E}_m\,\underline{\Gamma}_{N,N}\underline{V}_r\underline{\Sigma}_r^{-1/2} \qquad [5.1.16b]$$

$$\underline{A}_r = \underline{\Sigma}_r^{-1/2}\underline{Z}_r^T(\underline{\Gamma}^{\uparrow}_{N,N})\underline{V}_r\underline{\Sigma}_r^{-1/2} \qquad [5.1.16c]$$

where $\underline{A}_r$ is the unique rxr solution to

$$\underline{Q}_r \cdot \underline{A}_r \cdot \underline{P}_r = \underline{\Gamma}^{\uparrow}_{N,N} \qquad [5.1.17]$$

If $\sigma_i(\underline{\Gamma}_{N,N}) \neq 0$ for i>r+1, $\underline{A}_r$ will not be an exact solution to equation 5.1.15. The resulting realization will not be optimal because $\underline{\Gamma}_{r,r}$ is generally not a impulse hankel matrix. However, if $\sigma_r(\underline{\Gamma}_{N,N}) \gg \sigma_{r+1}(\underline{\Gamma}_{N,N})$, the realization $(\underline{A}_r,\underline{B}_r,\underline{C}_r)$ will adequately represent the process. $\underline{\Gamma}^{\uparrow}_{N,N}$ is used in the derivation of $\underline{A}_r$ instead of $\underline{\Gamma}_{N,N}$ so that the rank condition of the Ho-Kalman algorithm is not violated. The identification procedure will be very sensitive to noise because $\underline{\Gamma}_{N,N}$ is utilized twice in the algorithm: to obtain the singular values and singular vector matrices and to calculate $\underline{A}_r$. If r is chosen too large, there may be some small singular values which can decrease the robustness of the algorithm when they are inverted. Also, if r is chosen too small, there may not be enough state space dimensions to accurately approximate $\underline{A}$ (Zeigler, 1974).

Damen and Hajdasinski (1982) used the same method as Zeigler (1974) but used $\underline{\Gamma}_{r,r}$ and $\underline{\Gamma}^{\uparrow}_{r,r}$ to calculate $\underline{A}_r$, $\underline{B}_r$ and $\underline{C}_r$ instead of $\underline{\Gamma}_{N,N}$ and $\underline{\Gamma}^{\uparrow}_{N,N}$:

$$\underline{B}_r = \underline{\Sigma}_r^{-1/2}\underline{Z}_r^T\underline{\Gamma}_{r,r}\underline{E}_p^T \qquad [5.1.18a]$$

$$\underline{C}_r = \underline{E}_m\,\underline{\Gamma}_{r,r}\underline{V}_r\underline{\Sigma}_r^{-1/2} \qquad [5.1.18b]$$

$$\underline{A}_r = \underline{\Sigma}_r^{-1/2}\underline{Z}_r^{T}(\underline{\Gamma}^{\uparrow}_{r,r})\underline{V}_r\underline{\Sigma}_r^{-1/2} \qquad [5.1.18c]$$

where $\underline{A}_r$ is the unique rxr solution to

$$\underline{Q}_r \cdot \underline{A}_r \cdot \underline{P}_r = \underline{\Gamma}^{\uparrow}_{r,r} \qquad [5.1.19]$$

The calculation of $\underline{B}_r$ and $\underline{C}_r$ by equations 5.1.18a and 5.1.18b will be equivalent to that calculated from equations 5.1.15a and 5.1.15b because only the first (block) row and the first (block) column of the impulse hankel matrix are used to determine $\underline{B}_r$ and $\underline{C}_r$. However, $\underline{A}_r$ derived from equation 5.1.17c may be different from that obtained from equation 5.1.15a if the process is noisy or, r<n because the shifted truncated matrix is used instead of the shifted full matrix. This method meets the rank constraints of the Ho-Kalman algorithm but $\underline{\Gamma}_{r,r}$ and $\underline{\Gamma}^{\uparrow}_{r,r}$ does not retain hankel symmetry (equal blocks at cross-diagonals).

Kung (1978) also developed a similar modified Ho-Kalman algorithm to identify an approximate realization for stochastic systems similar to Zeigler's method. In Kung's algorithm, the minimal realization is obtained by shifting the factored matrices $\hat{\underline{Z}}_r$ or $\hat{\underline{V}}_r$ in equation 5.1.13 instead of the $\underline{\Gamma}_{N,N}$ or $\underline{\Gamma}_{r,r}$. Shifting of $\hat{\underline{Z}}_r$ up by one (block) row, $\hat{\underline{Z}}_r^{\uparrow}$, and shifting $\hat{\underline{V}}_r$ to the left by one (block) column, $\hat{\underline{V}}_r^{\leftarrow}$, is done by adding a row of zeros or a column of zeros to fill the matrices. Therefore, Kung assumes that h(i) = 0 when i>N. The r^th^ order approximate realization can be obtained from

$$\underline{A}_r = \hat{\underline{Z}}_r^{-1} \cdot \hat{\underline{Z}}_r^{\uparrow} = \hat{\underline{V}}_r^{\leftarrow} \cdot \hat{\underline{V}}_r^{-1}$$

$$= \underline{\Sigma}_r^{-1/2} \cdot \underline{Z}_r^{T} \cdot (\underline{Z}_r \cdot \underline{\Sigma}_r^{1/2})^{\uparrow} = (\underline{\Sigma}_r^{1/2} \cdot \underline{V}_r^{T})^{\leftarrow} \cdot \underline{V}_r \cdot \underline{\Sigma}_r^{-1/2} \qquad [5.1.20a]$$

$$\underline{B}_r = \underline{\Sigma}_r^{1/2} \cdot \underline{V}_r^{T} \qquad [5.1.20b]$$

$$\underline{C}_r = \underline{Z}_r \cdot \underline{\Sigma}_r^{1/2} \qquad [5.1.20c]$$

where $\underline{B}_r$ is the first r rows and the first (block) column of $\hat{\underline{V}}_r$ and $\underline{C}_r$ is the first r columns and the first (block) row of $\hat{\underline{Z}}_r$. Both the rank and hankel symmetry conditions of the Ho-Kalman algorithm are violated in Kung's method (Damen, 1982). Therefore, the longer the impulse sequence used (the larger the impulse hankel matrix), the more accurate Kung's algorithm will be since he assumes that $h_i = 0$ if $i>N$.

Hajdasinski and Damen (Damen, 1982) modified their previous method to meet the hankel rank and symmetry conditions of the Ho-Kalman algorithm by filtering $\underline{\Gamma}_{r,r}$. Given $\underline{\Gamma}_{r,r}$, which is obtained from the SVD of $\underline{\Gamma}_{N,N}$, a impulse hankel matrix $\hat{\underline{\Gamma}}_{r,r}$ is found from minimization of the spectral norm of $\underline{\Gamma}_{r,r}-\hat{\underline{\Gamma}}_{r,r}$ by

$$\frac{\partial \ \mathrm{tr}[(\underline{\Gamma}_{r,r}-\hat{\underline{\Gamma}}_{r,r})(\underline{\Gamma}_{r,r}-\underline{\Gamma}_{r,r})^T]}{\partial \hat{\underline{M}}(i+1)} = 0 \quad i=1,2,\ldots,2r-2 \qquad [5.1.21]$$

for $(i = 1,2,\ldots,2r-2)$ where $\mathrm{rank}(\hat{\underline{\Gamma}}_{r,r}) = \mathrm{rank}(\underline{\Gamma}_{r,r})$. $\hat{\underline{M}}(i)$ are $(p\times q)$ blocks in $\hat{\underline{\Gamma}}_{r,r}$ of approximated impulse response parameters. The minimization procedure is effectively averaging along the cross-diagonal blocks. Although the resulting impulse hankel matrix $\hat{\underline{\Gamma}}_{r,r}$ is hankel symmetric, the rank may not equal r. If this is the case, SVD is applied to $\hat{\underline{\Gamma}}_{r,r}$ and the above procedure is iterated until the rank of $\hat{\underline{\Gamma}}_{r,r}$ is r. The resulting matrix will hopefully converge to the best least squares approximation of $\underline{\Gamma}_{N,N}$.

Eydgahi (1987) obtained partial realizations from the infinite impulse response using a modified impulse hankel matrix in a Ho-Kalman identification algorithm. An r^{th} rank impulse hankel matrix, $\underline{\Gamma}_{N,N}$ is obtained from the input/output data of the process such that rank $(\underline{\Gamma}_{N,N})$ = rank $(\underline{\Gamma}_{N+1,N+1})$. A polynomial impulse hankel matrix, $\underline{\Gamma}_{N,N}(a)$ is formed which includes an arbitrary parameter 'a'. Using the original Ho-Kalman algorithm, $\underline{\Gamma}_{N,N}(a)$ is factored such that

$$\underline{R}(a)\cdot\underline{\Gamma}_{N,N}(a)\cdot\underline{S}(a) = \begin{bmatrix} \underline{I}_r & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} = \underline{U}^T\cdot\underline{U} \qquad [5.1.22]$$

where $\underline{I}_r$ is an rxr identity matrix, r is the rank of $\Gamma_{N,N}$ $(r \leq N)$ and

$$\underline{U} = [\ \underline{I}_r \quad \underline{0}\ ] \qquad [5.1.23]$$

The resulting rth order minimal realization is given by

$$\underline{A}(a) = \underline{U}\cdot[\underline{R}(a)\cdot\underline{\Gamma}^{\uparrow}_{N,N}(a)\cdot\underline{S}(a)]\cdot\underline{U}^T \qquad [5.1.24a]$$

$$\underline{B}(a) = \underline{U}\cdot\underline{R}(a)\cdot\underline{\Gamma}_{N,N}(a)\cdot\underline{E}_p^T \qquad [5.1.24b]$$

$$\underline{C}(a) = \underline{E}_m\cdot\underline{\Gamma}_{N,N}(a)\cdot\underline{S}(a)\cdot\underline{U}^T \qquad [5.1.24c]$$

By choosing an appropriate value of 'a', a stable realization can always be obtained if the original model is stable. However, this identification method involves manipulation of polynomial matrices.

Silverman (1980) and Maciejowski (1985) showed that the minimal partial realizations for a deterministic system obtained from Kung's algorithm and the other variations on the Ho-Kalman algorithm are approximately internally balanced. An exact balanced realization would be obtained only if the approximation to the impulse hankel matrix, $\underline{\Gamma}_{r,r}$, used in the identification algorithm is a impulse hankel matrix. Of the algorithms investigated, only the modified algorithm of Hajdasinski and Damen (Damen, 1982) exhibits this property. From Moore's work (1979a, 1981), it is known that the controllability and observability Grammians of the system can be factored in three ways:

1. input normalized: $\underline{W}_c^2 = \underline{P}^T\underline{P} = \underline{I}$ and $\underline{W}_o^2 = \underline{Q}\underline{Q}^T = \underline{\Sigma}^2$
2. output normalized: $\underline{W}_c^2 = \underline{P}^T\underline{P} = \underline{\Sigma}^2$ and $\underline{W}_o^2 = \underline{Q}\underline{Q}^T = \underline{I}$
3. internally balanced: $\underline{W}_c^2 = \underline{P}^T\underline{P} = \underline{\Sigma}$ and $\underline{W}_o^2 = \underline{Q}\underline{Q}^T = \underline{\Sigma}$

where $\underline{\Sigma}^2$ is the diagonal matrix of the square of the second order modes of the system. If the impulse hankel matrix is given by equation 5.1.5, then

$$\underline{\Gamma}_{N,N} = [\underline{Z}_r \;\; \underline{Z}_2] \cdot \begin{bmatrix} \underline{\Sigma}_r & 0 \\ 0 & \underline{\Sigma}_2 \end{bmatrix} \cdot \begin{bmatrix} \underline{V}_r^T \\ \underline{V}_2^T \end{bmatrix} = \underline{Q} \cdot \underline{P}$$

$$= \underline{Z}_r \Sigma_r \underline{V}_r^T + \underline{Z}_2 \Sigma_2 \underline{V}_2^T = \underline{Q}_r \cdot \underline{P}_r + \underline{Q}_2 \cdot \underline{P}_2 \qquad [5.1.25]$$

Assuming $\underline{\Sigma}_2 \ll \underline{\Sigma}_r$, $\underline{\Gamma}_{N,N}$ can be approximated by

$$\underline{\Gamma}_{r,r} = \underline{Z}_r \Sigma_r \underline{V}_r^T = \underline{Q}_r \cdot \underline{P}_r \qquad [5.1.26]$$

However, if $\underline{\Gamma}_{r,r}$ is not a hankel matrix, $\underline{Q}_r$ and $\underline{P}_r$ will only approximate the controllability and observability matrices of the system, $\underline{Q}$ and $\underline{P}$. The controllability and observability Grammians of the identified systems can be approximated by

$$\underline{W}_c^2 \cong \underline{P}_r \underline{P}_r^T = \underline{\Sigma}_r^{1/2} \underline{V}_r^T \underline{V}_r \underline{\Sigma}_r^{1/2} = \underline{\Sigma}_r \cong \tilde{\underline{\Sigma}} \qquad [5.1.27a]$$

$$\underline{W}_o^2 \cong \underline{Q}_r^T \underline{Q}_r = \underline{\Sigma}_r^{1/2} \underline{Z}_r^T \underline{Z}_r \underline{\Sigma}_r^{1/2} = \underline{\Sigma}_r \cong \tilde{\underline{\Sigma}} \qquad [5.1.27b]$$

In this case, the diagonal elements of $\underline{W}_c^2$ and $\underline{W}_o^2$ will only approximate the second order modes if $\underline{\Gamma}_{r,r}$ is not a hankel matrix. Since the singular values of the controllability and observability Grammians are approximately equal to the second order modes of the system, the resulting realizations are approximately internally balanced. The closer $\underline{Q}_r \underline{P}_r$ is to a hankel matrix, the more balanced the realization.

Example 5.1: A model for the discrete system (example 3.1 with a sampling time of 0.1) with state matrices given by:

$$\underline{A} = \begin{bmatrix} 3.320 & -4.312 & 2.594 & -0.607 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \underline{C}^T = \begin{bmatrix} 0.065 \\ 0.025 \\ -0.071 \\ 0.020 \end{bmatrix}$$

[5.1.28]

can be identified from the impulse response using the Ho-kalman algorithms. The program hokalman.m written in matlab (in Appendix C) was used to simulate the process and identify a model from its impulse response. If 20 impulse response parameters are used in the identification, the 4th order realization obtained using Zeigler's algorithm (equations 5.1.16a, 5.1.16b, 5.1.16c) is:

$$\underline{A} = \begin{bmatrix} 0.940 & -0.184 & 0 & -0.016 \\ 0.189 & 0.719 & -0.352 & -0.007 \\ 0 & 0.352 & 0.850 & -0.177 \\ 0.016 & -0.007 & 0.177 & 0.811 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} 0.739 \\ -0.751 \\ 0.291 \\ -0.049 \end{bmatrix} \quad \underline{C}^T = \begin{bmatrix} 0.739 \\ 0.751 \\ 0.291 \\ 0.049 \end{bmatrix}$$

[5.1.29]

The controllability and observability Grammians for the realization in equation 5.1.29 are given by

$$\underline{W}_c^2 = \begin{bmatrix} 2.746 & -1.075 & -0.162 & 0.033 \\ -1.075 & 0.864 & -0.197 & 0.016 \\ -0.162 & -0.197 & 0.177 & -0.025 \\ 0.033 & 0.016 & -0.025 & 0.005 \end{bmatrix} \quad \underline{W}_o^2 = \begin{bmatrix} 2.779 & 1.092 & -0.170 & -0.034 \\ -1.092 & 0.869 & 0.194 & 0.015 \\ -0.170 & 0.194 & 0.179 & 0.025 \\ -0.034 & 0.015 & 0.025 & 0.005 \end{bmatrix}$$

[5.1.30a]

with singular values

$$\{\sigma_i(\underline{W}_c^2) = \sigma_i(\underline{W}_o^2)\text{: } 3.260,\ 0.549,\ 0.012,\ 0\} \qquad [5.1.30b]$$

An equivalent balanced realization for the system given in equation 5.1.28 has singular values of

$$\{\sigma_i(\underline{W}_c^2) = \sigma_i(\underline{W}_o^2)\text{: } 2.448,\ 0.380,\ 0.017,\ 0\} \qquad [5.1.31]$$

Therefore, a second order model will approximate the system with an error of ≤ 0.012. The identified model is approximately a balanced realization since the controllability and observability Grammians are equal. However, the realization is not optimally balanced because the Grammians are not diagonal. The realization obtained from Kung's algorithm (5.1.20a, 5.1.20b, and 5.1.20c) is different and is given by:

$$\underline{A} = \begin{bmatrix} 0.919 & -0.179 & -0.031 & -0.007 \\ 0.236 & 0.698 & -0.285 & -0.027 \\ -0.232 & 0.459 & 0.513 & -0.078 \\ 1.523 & -0.705 & 2.370 & 0.169 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 0.739 \\ -0.751 \\ 0.291 \\ -0.049 \end{bmatrix} \qquad \underline{C}^T = \begin{bmatrix} 0.739 \\ 0.751 \\ 0.291 \\ 0.049 \end{bmatrix}$$

[5.1.32]

The singular values for the controllability and observability Grammians for the realization in equation 5.1.32 are:

$$\{\sigma_i(\underline{W}_c^2): 8.505,\ 1.659,\ 1.102,\ 0.014\} \qquad [5.1.33a]$$

$$\{\sigma_i(\underline{W}_o^2): 3.225,\ 0.562,\ 0.013,\ 0.002) \qquad [5.1.33b]$$

This model is not balanced.

Step responses and frequency responses for the identified models and the original model are shown in Figure 5.1a and Figure 5.1b, respectively. The 4th order realizations identified from 20 impulse response parameters using Hajdasinski's algorithms (equations 5.1.18a, 5.1.18b, and 5.1.18c) and Zeigler's method are equivalent to the model in equation 5.1.28. Hajdansinski's model and Zeigler's model should be the same since $\underline{\Gamma}_{r,r}$ is equal to $\Gamma_{n,n}$ and r=n. However, the realization identified using Kung's algorithm, equation 5.1.27, is not equivalent to the original model. The identified model captures information over a very narrow frequency range. The high frequency behaviour and the very low frequency behaviour of the process is not modelled accurately. In this case, Kung's algorithm assumes that the impulse response goes to zero after 20 time steps. However, the settling time for this system is approximately 50 time steps. Therefore, process information is neglected resulting in a poor model of the process. If 50 impulse response parameters are used in the identification of a 4th order realization, all four algorithms will produce equivalent state space models. Figure 5.2a and Figure 5.2b show the step responses and the frequency responses for these models.

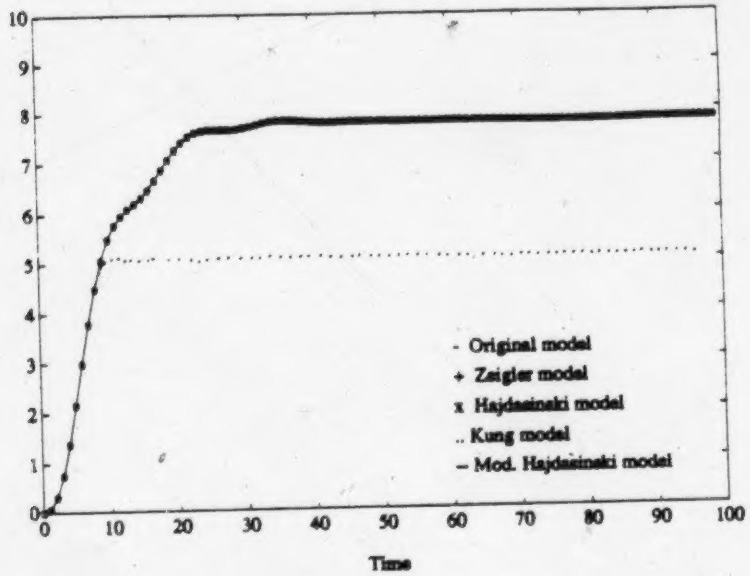

Figure 5.1a: Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters

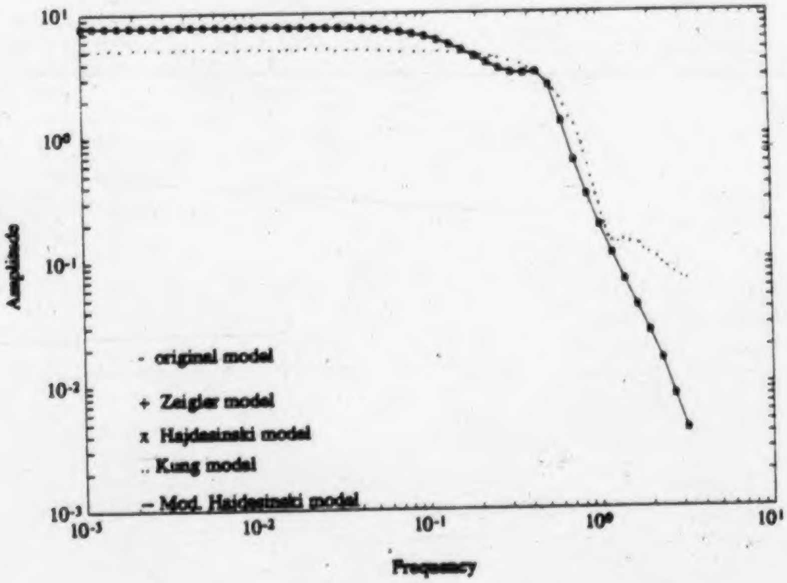

Figure 5.1b: Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters

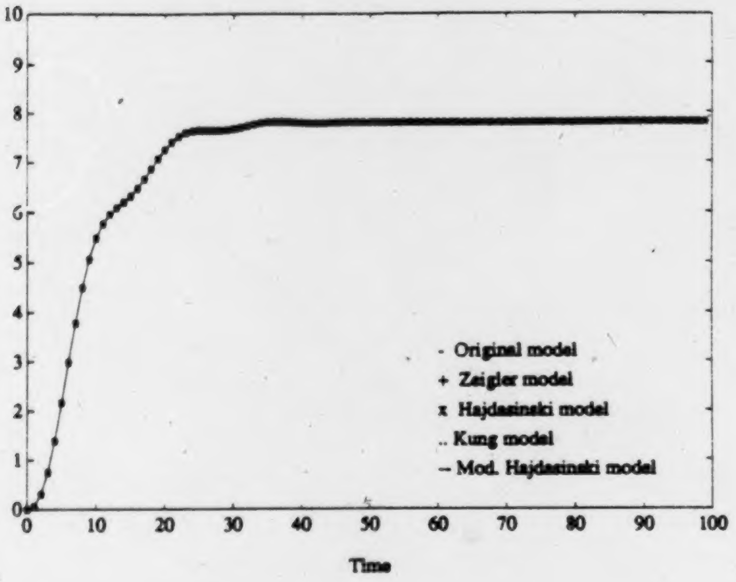

Figure 5.2a: Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters

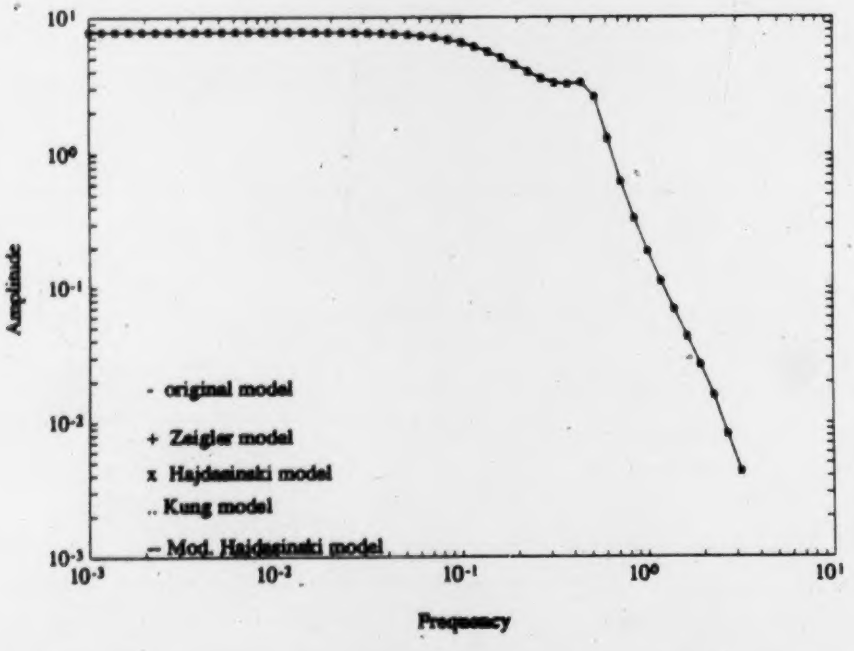

Figure 5.2b: Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters

As the number of impulse response parameters used in the identification algorithm's decrease, the steady state offset in the realization from Kung's algorithm increases. The identified models progressively ignore the high frequency and low frequency behaviour of the process. However, the models from the other three algorithm's still accurately describe the process. Figure 5.3a and Figure 5.3b show the step responses and the frequency responses for the four models identified using 10 impulse response parameters.

Lower order models can be identified from the Ho-Kalman algorithms. Figure 5.4a and Figure 5.4b show the step responses and the frequency responses for 3rd order models identified for the system in equation 5.1.28 using 50 impulse response parameters. All the realizations obtained from the Ho-Kalman algorithms exhibit a steady state offset from the original model. The identified models captured information over a wide frequency range. As the number of impulse response parameters used in the algorithm increases above 50, the steady state offset between the identified model and the original model remains constant. There is no new information contained within the impulse response sequence after 50 time steps. As the number of impulse response parameters used in the identification decreases below 50, the steady state offset increases. The model identified from Kung's algorithm exhibits significantly more offset than the models identified from the other three algorithms. The frequency range over which the models are valid decreases. Neglecting the impulse response parameters corresponds to neglecting the low frequency modes. Figure 5.5a and Figure 5.5b show the step response and the frequency responses for the 3rd order models obtained using 20 impulse response parameters. Figure 5.6a and Figure 5.6b show the step response and the frequency responses for the 3rd order models obtained

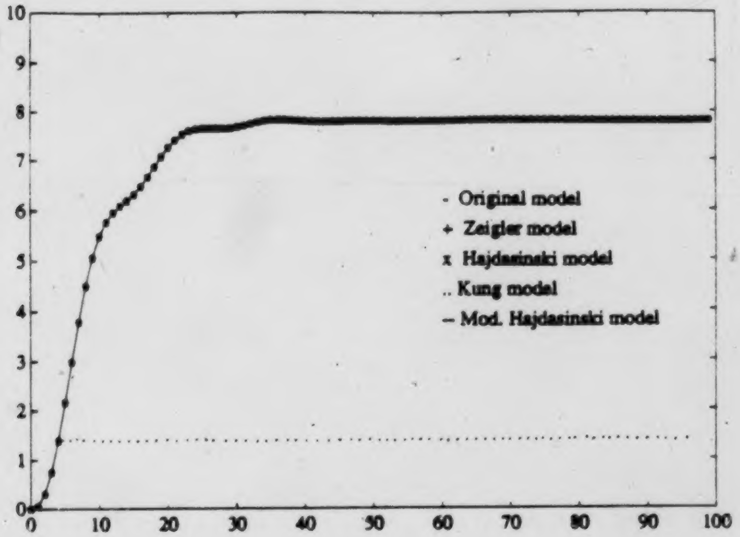

Figure 5.3a: Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 10 Impulse Response Parameters

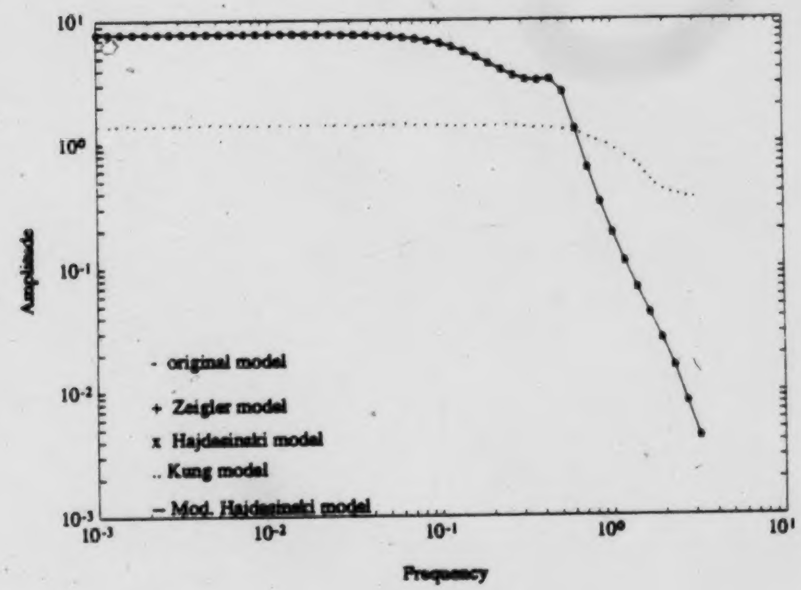

Figure 5.3b: Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 10 Impulse Response Parameters

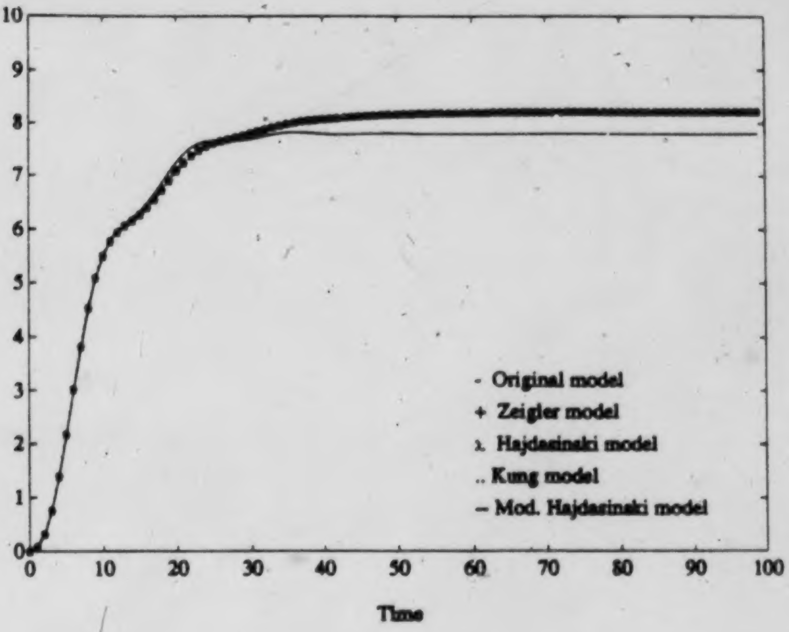

Figure 5.4a: Step Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters

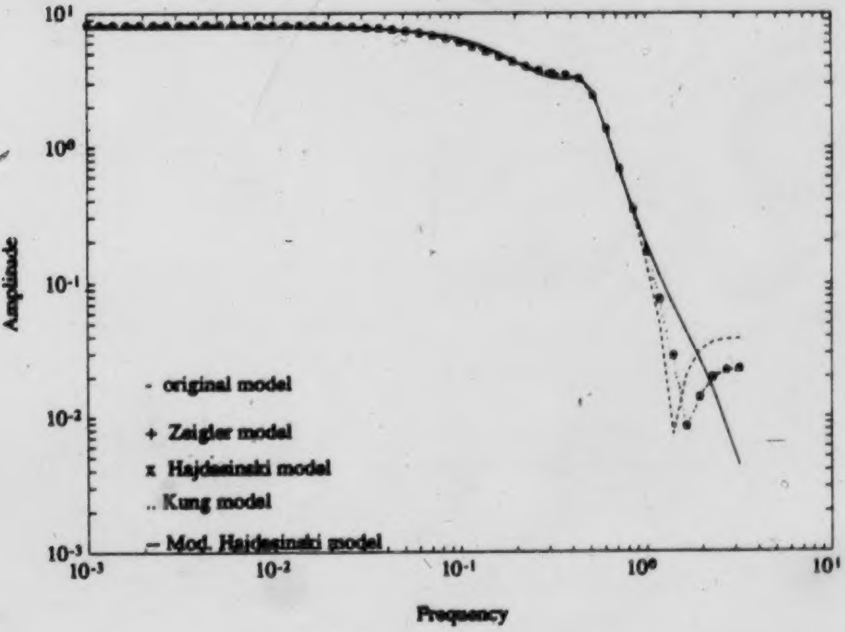

Figure 5.4b: Frequency Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters

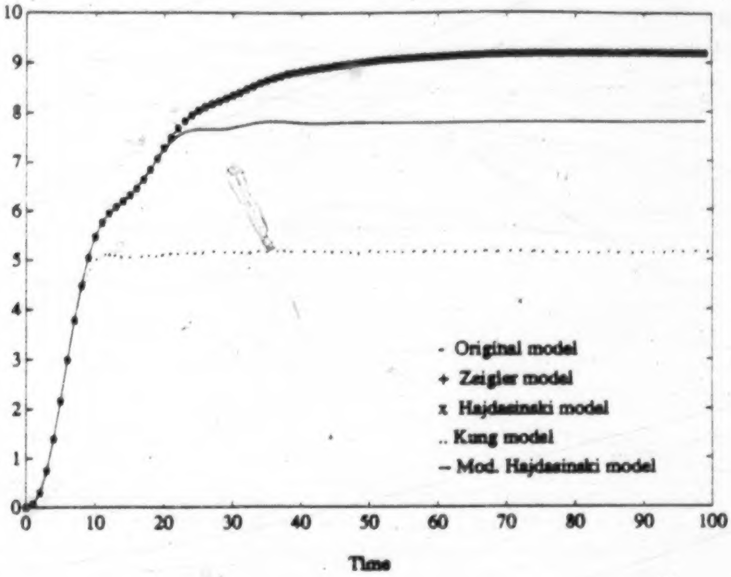

Figure 5.5a: Step Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters

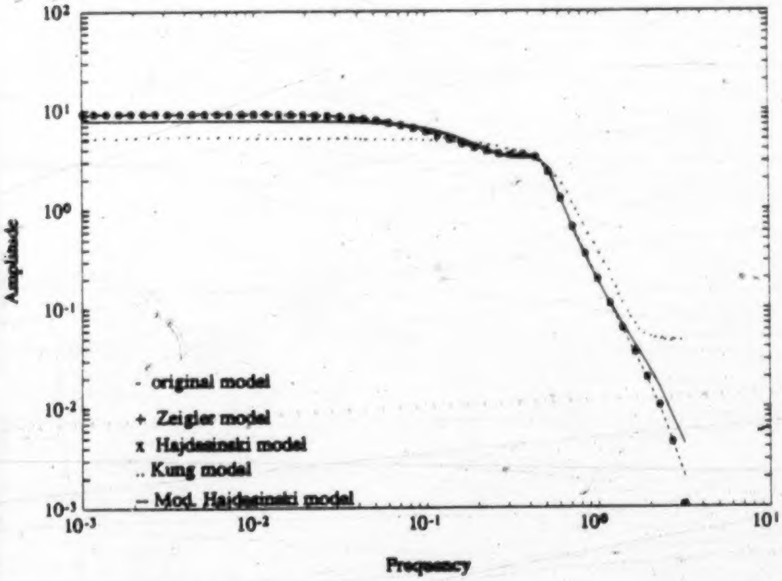

Figure 5.5b: Frequency Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters

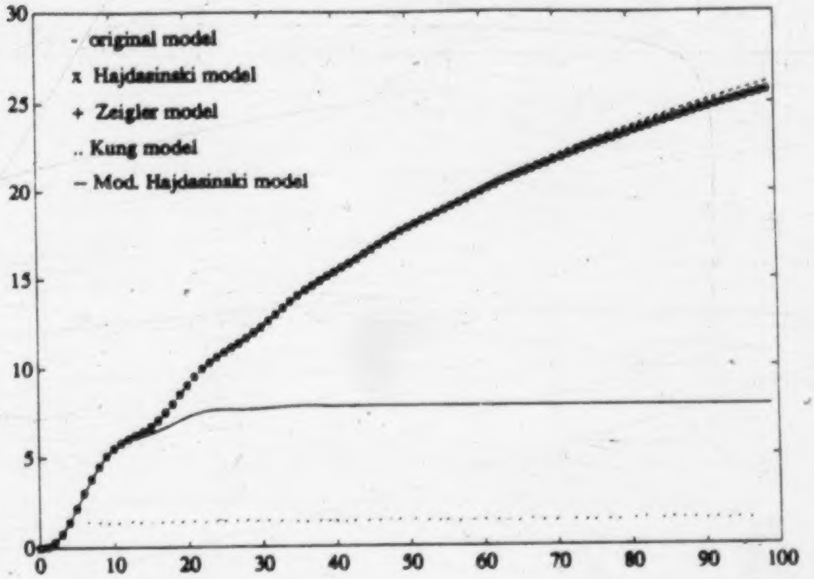

Figure 5.6a: Step Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 10 Impulse Response Parameters

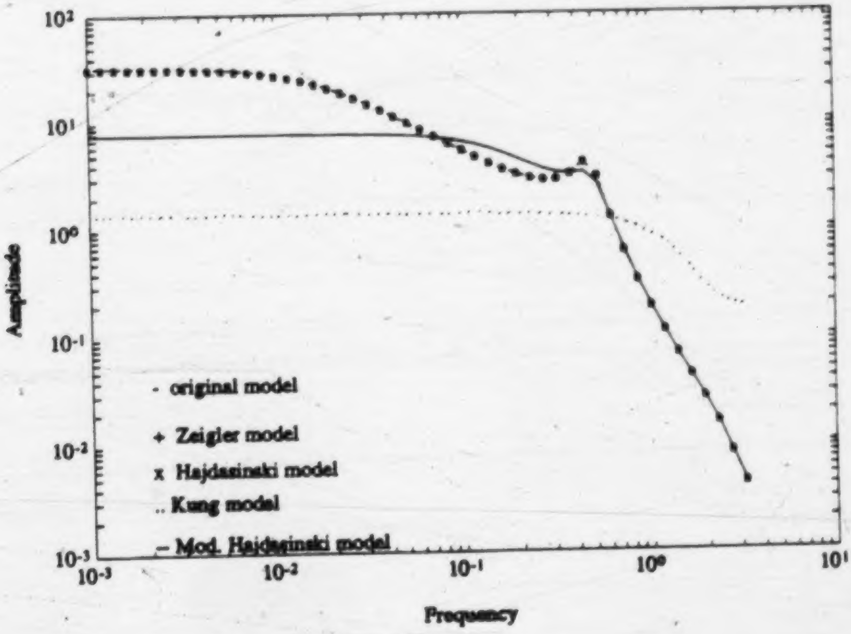

Figure 5.6b: Frequency Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 10 Impulse Response Parameters

using 10 impulse response parameters.

As the order of the model decreases, the Ho-Kalman algorithms neglect the high frequency modes of the process. If 50 impulse response parameters are used to identify a 2nd order model, the resulting realizations all exhibit steady state offset as shown in Figure 5.7a and Figure 5.7b. Figure 5.8a and Figure 5.8b show the step responses and the frequency responses for the 2nd order models obtained from 20 impulse response parameters. As with the 3rd order models, as the number of impulse response parameters used in the identification algorithm decreases, the lower frequency modes of the system are neglected.

Example 5.2: If noise is added to the system given in example 5.1, the identified models obtained using the Ho-Kalman algorithms can be different. If white noise with a variance of 0.1 is added to the system, the 4th order model obtained from Hajdasinski's algorithm and Zeigler's algorithm using 50 impulse response parameters are the same since $\underline{\Gamma}_{r,r}$ equals $\underline{\Gamma}_{n,n}$ and are given by:

$$\underline{A} = \begin{bmatrix} 0.935 & -0.164 & 0.003 & -0.003 \\ 0.164 & 0.621 & 0.236 & 0.048 \\ -0.003 & 0.236 & -0.817 & -0.509 \\ -0.002 & -0.043 & 0.509 & -0.856 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 0.757 \\ -0.803 \\ 0.103 \\ -0.011 \end{bmatrix} \qquad \underline{C}^T = \begin{bmatrix} 0.757 \\ 0.803 \\ -0.103 \\ -0.011 \end{bmatrix} \tag{5.1.34}$$

The 4th order model identified using Kung's algorithm with noise (variance 0.1) added to the system is given by

$$\underline{A} = \begin{bmatrix} 0.935 & -0.165 & 0.005 & -0.001 \\ 0.162 & 0.618 & 0.243 & 0.052 \\ -0.019 & 0.208 & -0.736 & -0.473 \\ -0.001 & -0.046 & 0.502 & -0.859 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 0.757 \\ -0.803 \\ 0.103 \\ -0.011 \end{bmatrix} \qquad \underline{C}^T = \begin{bmatrix} 0.757 \\ 0.803 \\ -0.103 \\ -0.011 \end{bmatrix} \tag{5.1.35}$$

The 4th order model identified using Hajdasinski's modified algorithm with

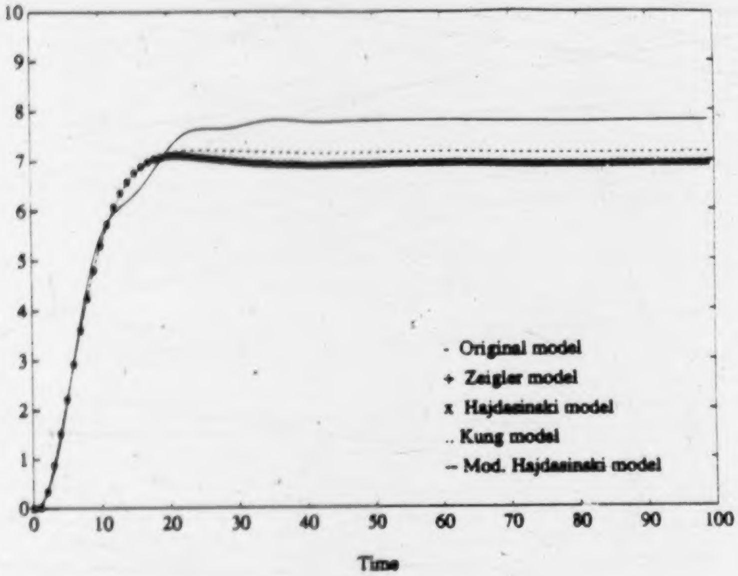

Figure 5.7a: Step Response for 2^{nd} Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters

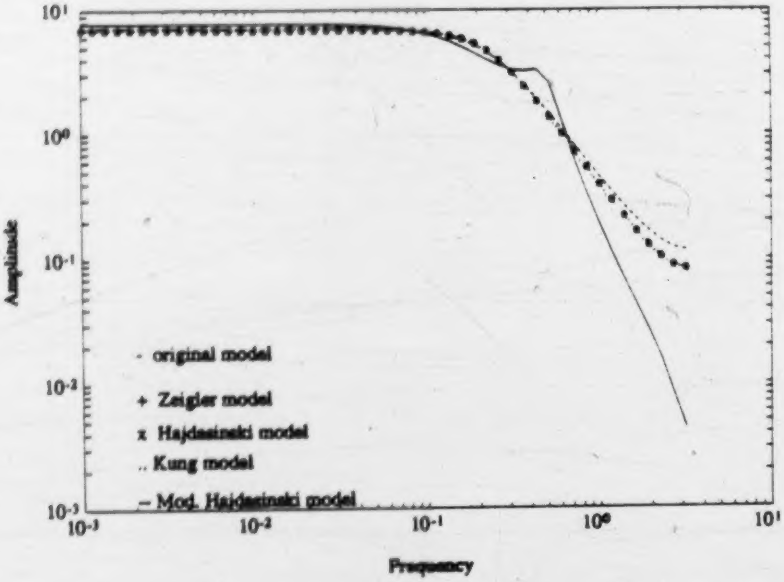

Figure 5.7b: Frequency Response for 2^{nd} Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters

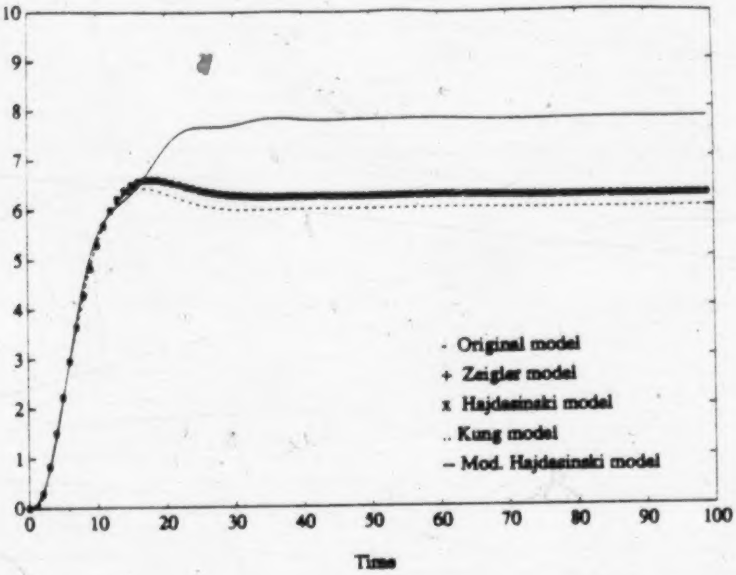

Figure 5.8a: Step Response for 2nd Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters

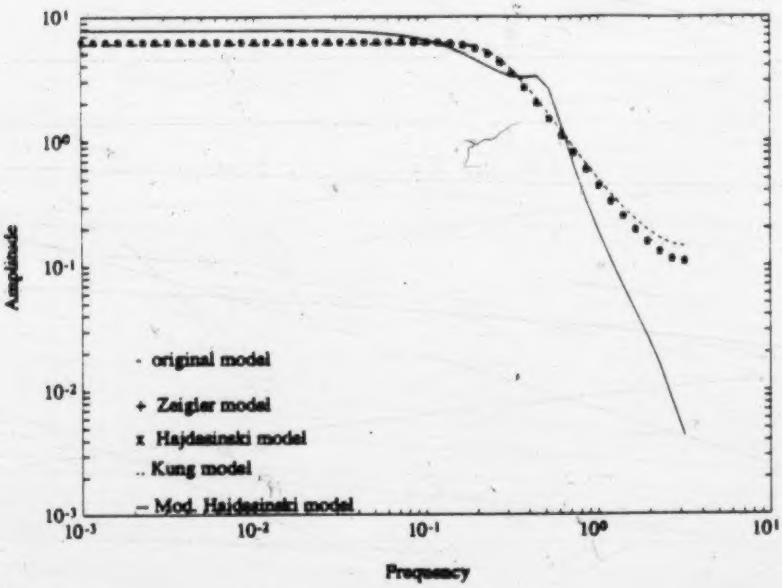

Figure 5.8b: Frequency Response for 2nd Order Models Identified from Ho-Kalman Algorithms Using 20 Impulse Response Parameters

noise (variance 0.1) added to the system is given by

$$\underline{A} = \begin{bmatrix} 0.935 & -0.164 & -0.017 & -0.001 \\ 0.164 & 0.525 & 0.380 & 0.123 \\ 0.017 & 0.380 & -0.748 & -0.507 \\ 0.001 & -0.123 & 0.507 & -0.860 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 0.759 \\ -0.828 \\ 0.001 \\ -0.028 \end{bmatrix} \qquad \underline{C}^T = \begin{bmatrix} 0.759 \\ 0.828 \\ -0.009 \\ -0.028 \end{bmatrix} \tag{5.1.36}$$

Figure 5.9a and Figure 5.9b show the step responses and the frequency responses for the four realizations. All the realizations exhibit a steady state offset. The high frequency modes of the system are not identified. Figure 5.10a and 5.10b show the step response and the frequency response for a 3rd order model identified using 50 impulse response parameters with noise of variance ∓ 0.1 added to the system. In this case, the steady state offset decreases. The identified models capture a wider range of frequency information neglecting only the high frequency information. The modified algorithm of Hajdasinski produces a better model than the other three algorithms since it accurately models more of the frequency response. Figure 5.11a and Figure 5.11b show the step response and frequency response for a 4th order model identified using 50 impulse parameters with noise of variance ∓0.05 added to the system. As the noise in the system decreases, the steady state offset also decreases. The identified models capture more of the frequency information with the modified Hajdasinski algorithm produces the best model.

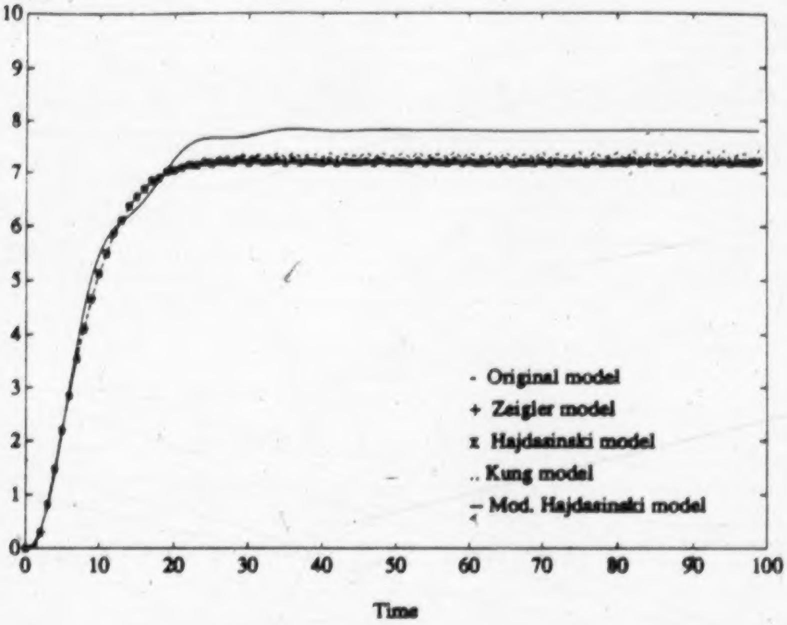

Figure 5.9a: Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters with Noise (Variance ∓0.1)

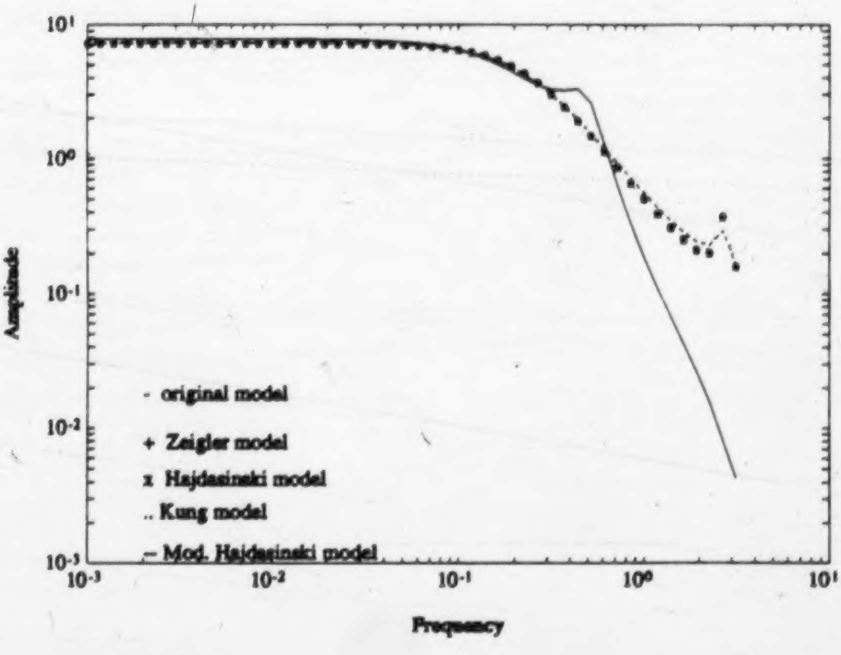

Figure 5.9b: Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters (Variance ∓0.1)

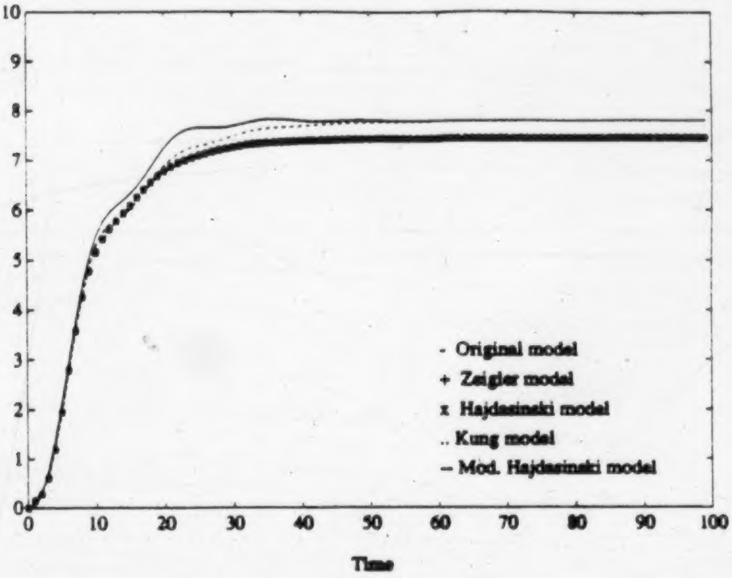

Figure 5.10a: Step Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters with Noise (Variance ∓0.1)

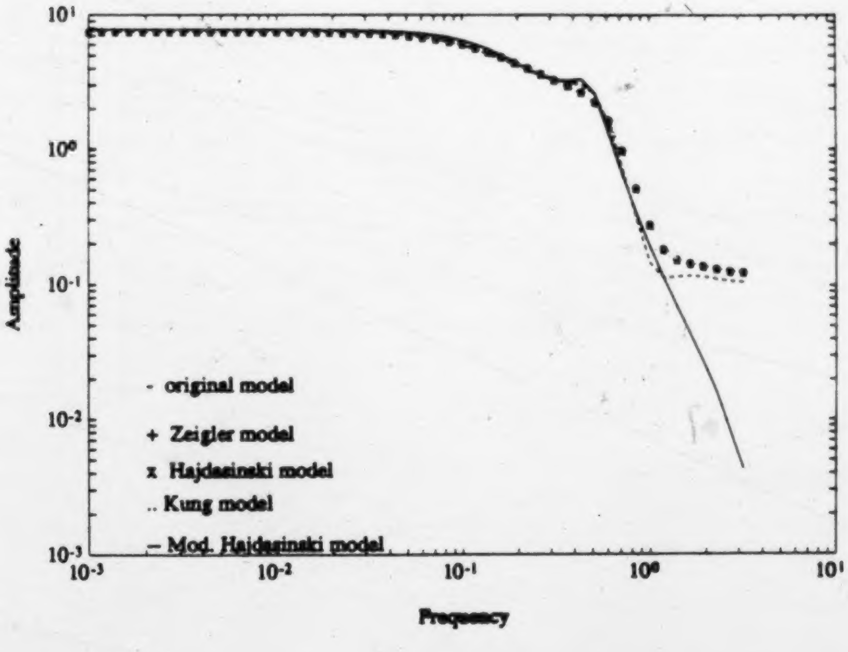

Figure 5.10b: Frequency Response for 3rd Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters (Variance ∓0.1)

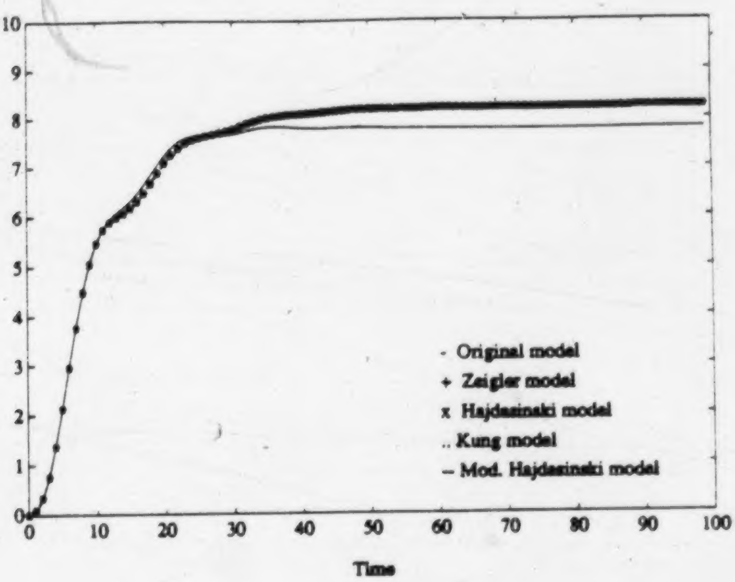

Figure 5.11a: Step Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters with Noise (Variance ∓0.05)

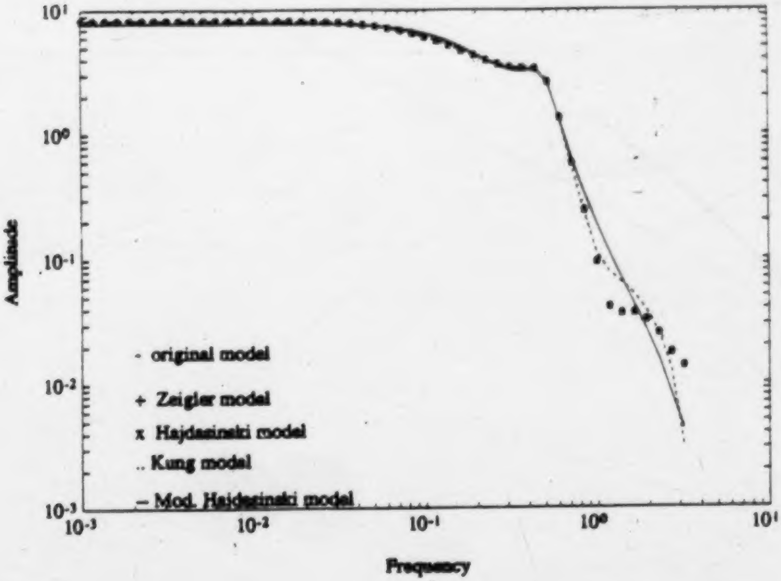

Figure 5.11b: Frequency Response for 4th Order Models Identified from Ho-Kalman Algorithms Using 50 Impulse Response Parameters (Variance ∓0.05)

5.2 Model Order Reduction Techniques Using SVD

Over the years, there have been several techniques developed for reducing the order of process models in the form of transfer functions and state space realizations. Chen (1974) and Tewari (1987) present an overview of conventional approaches which reduce the order of state space realizations. Bosley (1972) presents an overview of methods for deriving SISO transfer function models from high order state variable models.

Numerical methods have been used to fit a simple transfer function (first or second order) or matrix state equations to time response data or frequency response data of the process (curve fitting). A model is obtained which approximates the step response curves or impulse response curves by minimizing an error function over a finite interval (Anderson, 1967, Bosley 1972, Tewari, 1987). However, these methods retain many pitfalls of the numerical methods employed such as singularity. If frequency data is used, the low frequency range usually results in a poor fit. In addition, noise can affect the identified model and decrease its accuracy since the model is obtained from experimentally determined data.

Chen (1968) used continued fraction expansion of a high order transfer function model of the original system to reduce the model order. The high order model is expanded into a continued fraction and truncated. The resulting reduced order model will retain the dominant poles of the system. Although the method is simple, the algorithm may not converge and the reduced order models may be unstable even though the original model was stable (Eydgahi, 1987).

Several techniques have been developed to reduce the order of state space realizations which utilize physical reasoning to assess the relative importance of portions of the state space. Areas deemed to be unimportant are either ignored or approximated. Consequently, these types of model reduction schemes can be very subjective. Conventional modal reduction techniques eliminate the effects of stable modes (poles) and retain the dominant or slow modes (Litz, 1981) by neglecting states which correspond to large eigenvalues. State variables which strongly correspond to dominant eigenvalues are chosen to be state vector components of reduced order models. However, these methods can be used only if the original set of eigenvalues can be divided into two classes such that the effects of the undesired eigenvalues are important only at the beginning of the transient responses (Eydgahi, 1987). Davison (1966) neglected the eigenvalues which were farthest from the origin since these modes exhibited a negligible effect on the total system response.

Similarly, Moore (1979a, 1981), Silverman (1980), Glover (1984), Shokoohi (1987a), Tombs (1987) and several others have used principal component analysis to obtain asymptotically stable reduced order models from internally balanced state space models of asymptotically stable systems by neglecting a subset of the second order modes of the system. The model reduction techniques which utilize internally balanced state space models rely on the fact that the states are ordered from the most controllable/observable to the least controllable/observable to ensure that the "error" in the reduced order model is minimal. Weak subsystems (the least observable/controllable states) which contribute insignificantly to the input-output behaviour of the process are eliminated. The dominant subsystems (the most controllable/observable states) of the state space are retained. The use of internally balanced state space models and SVD results in a order reduction method which is numerically robust. In addition, the method can be applied to time variant systems as well as time invariant systems.

Retention of dominant characteristics and continued fraction expansion methods require a high order transfer function model or state space realization which accurately represents the process prior to performing the model reduction to obtain a "good" reduced order model. In a practical application, these high order models may not be available. In addition, the accuracy of models derived from experimental data is dependent on the signal/noise ratio and the occurrence of any undetermined disturbances during the identification. To overcome these difficulties, Wahlberg (1986) identified a high order (N^{th} degree) ARX model

$$A(q^{-1})y(t) = B(q^{-1})u(t) + e(t) \qquad [5.2.1]$$

from the impulse response of the process using a batch least squares identification technique based on N observations:

$$y(t)=\sum_{i=1}^{N} h_i q^{-i} u(t)=H(q^{-1})u(t) \qquad [5.2.2]$$

As N approaches infinity, the N^{th} order ARX model will approach the true model. The N^{th} order model was reduced to give a model of a reasonable order which would characterize the process. Wahlberg (1986) investigated three different model reduction techniques:

1. weighted L^2 model reduction which utilized a numerical non-linear optimization algorithm. The following optimization criteria is minimized with respect to ρ (parameters for the reduced order model) giving the minimum variance of the estimates:

$$J = \int_{-\pi}^{\pi} \left|\underline{G}(z,\theta_N)-\underline{G}(z,\rho)\right|^2 Q(\omega)\,d\omega \qquad [5.2.3]$$

where $z=e^{i\omega}$, θ_N is the parameter vector for the N order model, $\underline{G}(z,\rho)$ is the set of lower order transfer functions of order η, and $Q(\omega)$ is the ratio of the input spectrum to the noise spectrum as a function of frequency (ω).

2. model reduction via frequency weighted internally balanced realizations. The N^th^ order model is transformed to an internally balanced realization (Moore, 1979a). An input and/or output weighting filter, $\underline{W}(z)$, is added to the transfer function.
3. frequency weighted optimal hankel norm model reduction. In this method, a frequency weighted hankel norm is minimized to find the reduced order model.

The weighted L^2 model reduction technique will give an optimal L^2 estimate of the lower order transfer function which will minimize the prediction error between the actual process response and the predicted process response over the set of transfer functions of order η. However, the estimate may be a local optimum and not necessarily a global optimum (Wahlberg, 1986). In addition, this technique involves the use of numerical nonlinear optimization algorithms and an estimate of the noise and input spectrum. On the other hand, model reduction via frequency weighted internally balanced realizations or optimal hankel norm techniques provide an explicit solution and are relatively easy to implement. Although the reduced order models are not optimal, they will adequately describe the process. The risk of obtaining a locally optimal estimate is eliminated (Wahlberg, 1986).

5.2.1 Model Reduction via Truncation of Internally Balanced Realizations

Let a nth order time invariant internally balanced state space realization, $(\underline{A},\underline{B},\underline{C},\underline{D})$, of a continuous, asymptotically stable system be given by

$$\begin{bmatrix} \dot{\underline{x}}_1(t) \\ \dot{\underline{x}}_2(t) \end{bmatrix} = \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} \\ \underline{A}_{21} & \underline{A}_{22} \end{bmatrix} \cdot \begin{bmatrix} \underline{x}_1(t) \\ \underline{x}_2(t) \end{bmatrix} + \begin{bmatrix} \underline{B}_1 \\ \underline{B}_2 \end{bmatrix} \cdot \underline{u}(t) \qquad [5.1.4a]$$

$$\underline{y}(t) = \begin{bmatrix} \underline{C}_1 & \underline{C}_2 \end{bmatrix} \cdot \begin{bmatrix} \underline{x}_1(t) \\ \underline{x}_2(t) \end{bmatrix} + \begin{bmatrix} \underline{D}_1 \\ \underline{D}_2 \end{bmatrix} \cdot \underline{u}(t) \qquad [5.2.4b]$$

where $\underline{x}_1(t) \in \mathbb{R}^{r\times 1}$ and $\underline{x}_2(t) \in \mathbb{R}^{(n-r)\times 1}$. The matrix $\underline{H}$ in equation 4.1.4, the controllability Grammian and the observability Grammian can be partitioned in a similar manner such that

$$\underline{H} = \underline{\Sigma}_o \underline{Z}_o \underline{Z}_c \underline{\Sigma}_c = \begin{bmatrix} \underline{Z}_1 & \underline{Z}_2 \end{bmatrix} \cdot \begin{bmatrix} \underline{\Sigma}_1^2 & \underline{0} \\ \underline{0} & \underline{\Sigma}_2^2 \end{bmatrix} \cdot \begin{bmatrix} \underline{V}_1^T \\ \underline{V}_2^T \end{bmatrix} \qquad [5.2.5]$$

$$\underline{W}_c^2 = \begin{bmatrix} \underline{Z}_{c1} & \underline{Z}_{c2} \end{bmatrix} \cdot \begin{bmatrix} \underline{\Sigma}_{c1}^2 & \underline{0} \\ \underline{0} & \underline{\Sigma}_{c2}^2 \end{bmatrix} \cdot \begin{bmatrix} \underline{Z}_{c1}^T \\ \underline{Z}_{c2}^T \end{bmatrix} \qquad [5.2.6]$$

$$\underline{W}_o^2 = \begin{bmatrix} \underline{Z}_{o1} & \underline{Z}_{o2} \end{bmatrix} \cdot \begin{bmatrix} \underline{\Sigma}_{o1}^2 & \underline{0} \\ \underline{0} & \underline{\Sigma}_{o2}^2 \end{bmatrix} \cdot \begin{bmatrix} \underline{Z}_{o1}^T \\ \underline{Z}_{o2}^T \end{bmatrix} \qquad [5.2.7]$$

where $\underline{\Sigma}_1^2 = \mathrm{diag}(\sigma_1^2,\ldots,\sigma_r^2)$ and $\underline{\Sigma}_2^2 = \mathrm{diag}(\sigma_{r+1}^2,\ldots,\sigma_n^2)$ The controllability and observability Grammians of $(\underline{A},\underline{B},\underline{C},\underline{D})$ will be given by

$$\underline{W}_c^2 = \underline{W}_o^2 = \underline{\Sigma}_h = \mathrm{diag}(\underline{\Sigma}_1,\underline{\Sigma}_2) \qquad [5.2.8]$$

since the realization is internally balanced. As discussed in chapter 4, the controllability and observability subspaces of a realization are spanned by the singular vectors of the Grammians and the axis lengths are the singular values of the Grammians.

The state vector in equation 5.2.4 can be expressed in terms of the column vectors of $\underline{Z}_{c1}$ where

$$\tilde{\underline{x}}(t) = \underline{Z}_{c1}^{T}\underline{x}(t) \qquad [5.2.9]$$

The resulting state space realization $(\tilde{\underline{A}},\tilde{\underline{B}},\tilde{\underline{C}},\tilde{\underline{D}})$ corresponding to equation 5.2.4 will be given by equations 3.1.26 where the linear transformation matrix $\underline{T}$ is equal to $\underline{Z}_{c1}$. Since $(\underline{A},\underline{B},\underline{C},\underline{D})$ is internally balanced,

$$\tilde{\underline{A}} = \underline{A}_{11} \qquad [5.2.10a]$$

$$\tilde{\underline{B}} = \underline{B}_{1} \qquad [5.2.10b]$$

$$\tilde{\underline{C}} = \underline{C}_{1} \qquad [5.2.10c]$$

$$\tilde{\underline{D}} = \underline{D}_{1} \qquad [5.2.10d]$$

If $\underline{\Sigma}_{c2}^{2}=0$, the columns of $\underline{Z}_{c1}$ $(=\underline{Z}_{o1})$ will span the controllable (observable) subspace. The uncontrollable (unobservable) subspace will be spanned by the vectors corresponding to the columns of $\underline{Z}_{c2}$ $(=\underline{Z}_{o2})$. In this case, $(\tilde{\underline{A}},\tilde{\underline{B}},\tilde{\underline{C}},\tilde{\underline{D}})$ will be expressed in terms of basis vectors of the controllable subspace, X_c. Since the controllable subspace is A-invariant and $\underline{Z}_{c2}^{T}\underline{x}(t) = \underline{0}$, $(\underline{A}_{11},\underline{B}_{1},\underline{C}_{1},\underline{D}_{1})$ will have the same transfer function as $(\underline{A},\underline{B},\underline{C},\underline{D})$ (Moore, 1979a).

If $\underline{\Sigma}_{c2}^{2}\neq\underline{0}$ and $\sigma_r \gg \sigma_{r+1}$, the components of the state vector in the directions of the columns of $\underline{Z}_{c2}$, $\underline{x}_2(t)$, will be relatively small compared to $\underline{x}_1(t)$. In this case, the columns of $\underline{Z}_{c1}$ $(\underline{Z}_{o1})$, the major axes of the controllability (observability) ellipsoid, will define the strongly controllable (observable) subsystem, X_s. The columns of $\underline{Z}_{c2}$ $(\underline{Z}_{o2})$, the minor axes of the controllability (observability) ellipsoid, will define the weakly controllable (observable) subsystem, X_w, which is perpendicular to the strong subsystem. If $\sigma_r \gg \sigma_{r+1} \simeq \underline{0}$, the axis corresponding to the columns of $\underline{Z}_{c2}$ form a degenerate portion of the subspace which does not contribute to the input/output behaviour of the process. In this case, $(\underline{\tilde{A}},\underline{\tilde{B}},\underline{\tilde{C}},\underline{\tilde{D}})$ will be expressed in terms of the basis vectors of the strongly controllable (observable) subspace. However, X_s is not A-invariant and $\underline{Z}_{c2}^{T}\underline{x}(t) \neq \underline{0}$. Therefore, $(\underline{\tilde{A}},\underline{\tilde{B}},\underline{\tilde{C}},\underline{\tilde{D}})$ will have a different transfer function than $(\underline{A},\underline{B},\underline{C},\underline{D})$ because a portion of the controllable (observable) subspace has been neglected (Moore, 1979a). If $\sigma_r \gg \sigma_{r+1}$, the portion of the controllable subspace and the observable subspace corresponding to $\underline{x}_2(t)$ will have small magnitudes relative to the portion of the controllable subspace and the observable subspace corresponding to $\underline{x}_1(t)$ (Moore, 1979a, Pernebo, 1982). Therefore, $\underline{x}_2(t)$ will be much less affected by an input than $\underline{x}_1(t)$. Similarly, $\underline{x}_1(t)$ will be affected by the output considerably more that $\underline{x}_2(t)$ (Pernebo, 1982, Moore, 1979a). Therefore, if $\sigma_r \gg \sigma_{r+1}$, the effect of $\underline{x}_2(t)$ on the input/output behaviour of the system is not significant relative to the effect of $\underline{x}_1(t)$, the subsystem $(\underline{A}_{11},\underline{B}_1,\underline{C}_1,\underline{D}_1)$ is said to be internally dominant (Moore, 1979a, 1981). Moore (1979a) showed that an internally dominant subsystem of order r will exist for an n^th order system if and only if

$$\left\{ \sum_{i=1}^{r} \sigma_i^2(\underline{H}) \right\}^{1/2} \gg \left\{ \sum_{i=r+1}^{n} \sigma_i^2(\underline{H}) \right\}^{1/2} \qquad \text{[5.2.1]]}$$

Therefore, if $(\underline{A}_{11},\underline{B}_1,\underline{C}_1,\underline{D}_1)$ corresponds to an internally dominant subsystem, the r^th order realization will approximate the n^th order process $(\underline{A},\underline{B},\underline{C},\underline{D})$ and retain the dominant input/output properties of the system (Moore, 1981). Glover (1984) and Wahlberg (1986) showed that the L^{∞} bound on the error between the full process model $(\underline{G}(s))$ and the truncated balanced realization $(\underline{G}(s))$ is given by

$$\sigma_{r+1} \leq \|E(s)\|_{\infty} = \|\underline{G}(s) - \underline{G}_r(s)\|_{\infty} \leq 2 \sum_{i=r+1}^{n} \sigma_i(\underline{H}) \qquad \text{[5.2.12]}$$

The hankel norm of the error is defined as (Glover, 1984):

$$\|\underline{G}(s)-\underline{G}_r(s)\|_H \leq 2(\sigma_{r+1}+\sigma_{r+2}+\ldots+\sigma_n) \qquad [5.2.13]$$

Therefore, truncated balanced realizations will produce good reduced order models. If σ_{r+1} tends towards σ_r, the poles of the truncation tend to the imaginary axis (Glover, 1984). The modes corresponding to these singular values will become uncontrollable and unobservable resulting in small errors in the input/output reduced order model. Therefore, it may not be necessary for $\sigma_{r+1} \ll \sigma_r$ to get a good reduced order model.

In a similar manner, a reduced order model for an input-normalized or output-normalized realization can be obtained. These reduced order models will be algebraically equivalent to the internally balanced models of the same order (Moore, 1979a, 1981).

Asymptotic stability of all subsystems $\{(\underline{A}_{ii},\underline{B}_i,\underline{C}_i,\underline{D}_i)$: $i=1,2\}$ of a continuous internally balanced realizations obtained using Moore's algorithm will be guaranteed only if $\underline{\Sigma}_h$ has distinct singular values (Pernebo, 1982). If $\underline{\Sigma}_1$ and $\underline{\Sigma}_2$ have diagonal entries in common, there is no guarantee that the reduced order models will be asymptotically stable. However, it may be possible to find an equivalent balanced realization such that every subsystem is asymptotically stable (Pernebo, 1982) since a balanced realization will be unique (within a sign transformation) only if the singular values are unique. Therefore, Moore's reduction algorithm will always give a reduced order model which is controllable and observable but not necessarily asymptotically stable.

Pernebo (1982) extended the results of Moore (1979a, 1981) to discrete time systems. The system matrices for the discrete, asymptotically stable internally balanced realization can be partitioned as in equation 5.3.1 (for the continuous time case). Reduced order models $(\underline{A}_{11},\underline{B}_1,\underline{C}_1,\underline{D}_1)$ which strip away the weak structure in the process impulse hankel matrix can be obtained from the internally balanced realizations using Moore's algorithm. Al-Saggaf (Theorem 1, 1987) showed that the L^{∞} bound on the truncation error for the balanced realizations is given by:

$$\left\| \underline{E}(z) \right\|_{\infty} = \left\| \underline{G}(z) - \underline{G}_r(z) \right\|_{\infty} \leq 2 \sum_{i=r+1}^{n} \sigma_i \qquad [5.2.14]$$

where the strict inequality holds if the singular values are unique ($\sigma_i \neq \sigma_{i+1}$ for all $i=r+1,\ldots,n$).

Unlike the continuous time case, a more general result exists concerning the stability of subsystems of the discrete time systems. Pernebo (1982) showed that every subsystem of the discrete time system is asymptotically stable if the original system is asymptotically stable and either the observability or the controllability Grammian is non-singular and diagonal. Therefore, if the system is internally balanced, all subsystems will always be asymptotically stable. If $\sigma_r > \sigma_{r+1}$, $(\underline{A}_{11}, \underline{B}_1, \underline{C}_1, \underline{D}_1)$ will be controllable and observable. However, the other subsystems are not necessarily controllable and observable (Pernebo, 1982). Therefore, the conditions for controllability and observability of the discrete reduced order models is a weaker result than for the continuous case.

Example 5.3: For the continuous system given in example 3.1, the 4th order internally balanced realization is given by

$$\underline{A} = \begin{bmatrix} -0.518 & 1.450 & 0.391 & -0.350 \\ -1.450 & -2.195 & -4.753 & 1.128 \\ 0.391 & 4.753 & -0.630 & 1.196 \\ 0.350 & 1.218 & -1.196 & -1.657 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} -0.773 \\ -0.805 \\ 0.337 \\ 0.252 \end{bmatrix} \qquad \underline{C}^T = \begin{bmatrix} -0.773 \\ 0.805 \\ 0{,}337 \\ -0.252 \end{bmatrix} \qquad [5.2.15]$$

The controllability Grammian is equal to the observability Grammian such that

$$\underline{W}_c^2 = \underline{W}_o^2 = \text{diag}(0.576,\ 0.148,\ 0.090,\ 0.019). \qquad [5.2.16]$$

The diagonal elements of the Grammians are the second order modes of the system. First, second and third order internally balanced models can be obtained by deleting the states corresponding to the smallest singular values. The reduced models will be given by:

r=1

$$\underline{A} = \begin{bmatrix} -0.518 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} -0.773 \end{bmatrix} \qquad \underline{C} = \begin{bmatrix} -0.773 \end{bmatrix} \qquad [5.2.17]$$

r=2

$$\underline{A} = \begin{bmatrix} -0.518 & 1.450 \\ -1.450 & -2.195 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} -0.773 \\ -0.805 \end{bmatrix} \qquad \underline{C}^T = \begin{bmatrix} -0.773 \\ 0.805 \end{bmatrix} \qquad [5.2.18]$$

r=3

$$\underline{A} = \begin{bmatrix} -0.518 & 1.450 & 0.391 \\ -1.450 & -2.195 & -4.753 \\ 0.391 & 4.753 & -0.630 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} -0.773 \\ -0.805 \\ 0.337 \end{bmatrix} \qquad \underline{C}^T = \begin{bmatrix} -0.773 \\ 0.805 \\ 0.337 \end{bmatrix} \qquad [5.2.19]$$

Figure 5.12a and Figure 5.12b show the step responses and the frequency responses for the three reduced order models and the original internally balanced state space model. Although the reduced order models appear to approximate the high frequency behaviour, the low frequency dynamics are not modeled as well resulting in a steady state offset. The steady state offset increases as the model order is reduced from the true model order. However, the steady state offset appears to be less than that obtained with the Ho-kalman methods described in example 5.1.

Example 5.4: A system is described by (Glover, 1984) the internally balanced realization is:

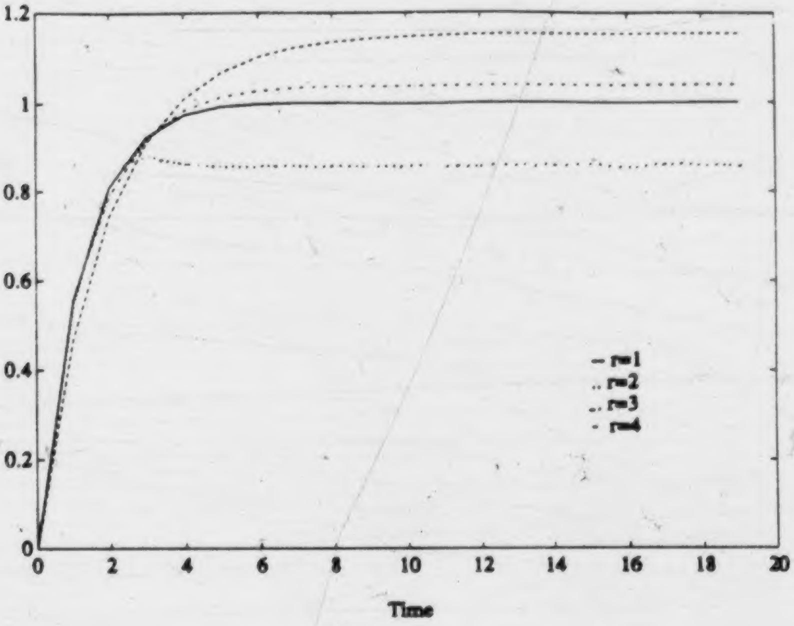

Figure 5.12a: Step Response for the Reduced Order Models Given in Example 5.3

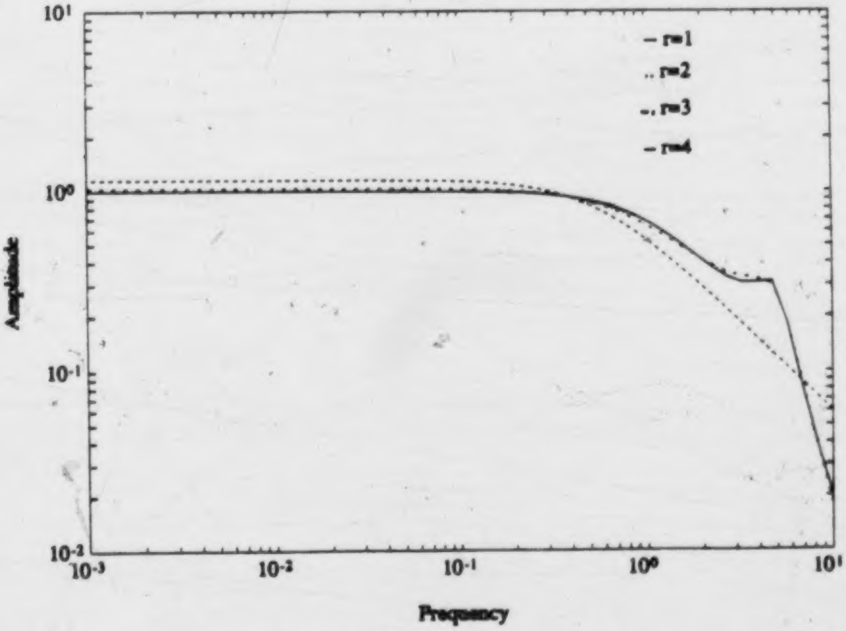

Figure 5.12b: Frequency Response for the Reduced Order Models Given in Example 5.3

$$\underline{A} = \begin{bmatrix} -9 & 4 & -4 \\ 4 & -2 & 4 \\ 4 & -4 & -1 \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} 6 \\ -2 \\ -1 \end{bmatrix} \qquad \underline{C}^T = \begin{bmatrix} 6 \\ -2 \\ 1 \end{bmatrix} \qquad [5.2.20]$$

The singular values for the system are {2.0, 1.0, 0.5}. Figure 5.13a and Figure 5.13b show the step responses and frequency responses for the internally balanced realizations of order 1, 2 and 3. As in example 5.3, there is steady state offset with all the reduced order models. The lower order models have larger offsets than the higher order models.

5.2.2 Frequency Weighted Truncated Internally Balanced Models

Al-Saggaf (1987) and Wahlberg (1986) used the model reduction method of Moore to reduce the order of frequency weighted internally balanced, input normalized and output normalized state space realizations. Frequency weighting can be used to make the reduction error approximately uniform over all frequencies or to minimize the reduction errors within a certain frequencyband. An input and/or output weighting filter, $\underline{W}_u(z)$ or $\underline{W}_y(z)$, is added to the process transfer function (matrix) as shown in Figure 5.14.

Let $(\underline{A}_w, \underline{B}_w, \underline{C}_w)$ be the minimal realization of the frequency weighting matrix, $\underline{W}_u(z)$, applied to the input signals where $\underline{A}_w \in \mathbb{R}^{nxn}, \underline{B}_w \in \mathbb{R}^{nxp}, \underline{C}_w \in \mathbb{R}^{mxn}$. Frequency weighted discrete time Lyapunov equations can be solved to find the discrete time frequency weighted controllability and observability Grammians

$$\underline{A}\underline{W}_c^2\underline{A}^T - \underline{W}_c^2 + \underline{X}\underline{B}_w\underline{B}_w^T\underline{X}^T = 0 \qquad [5.2.21a]$$

$$\underline{A}^T\underline{W}_o^2\underline{A} - \underline{W}_o^3 + \underline{C}^T\underline{C} = 0 \qquad [5.2.21b]$$

$$\underline{A}\underline{X} - \underline{X}\underline{A}_w + \underline{B}\underline{C}_w = 0 \qquad [5.2.21c]$$

If $(\underline{A}, \underline{B}, \underline{C}, \underline{D})$ is asymptotically stable and observable, the controllability

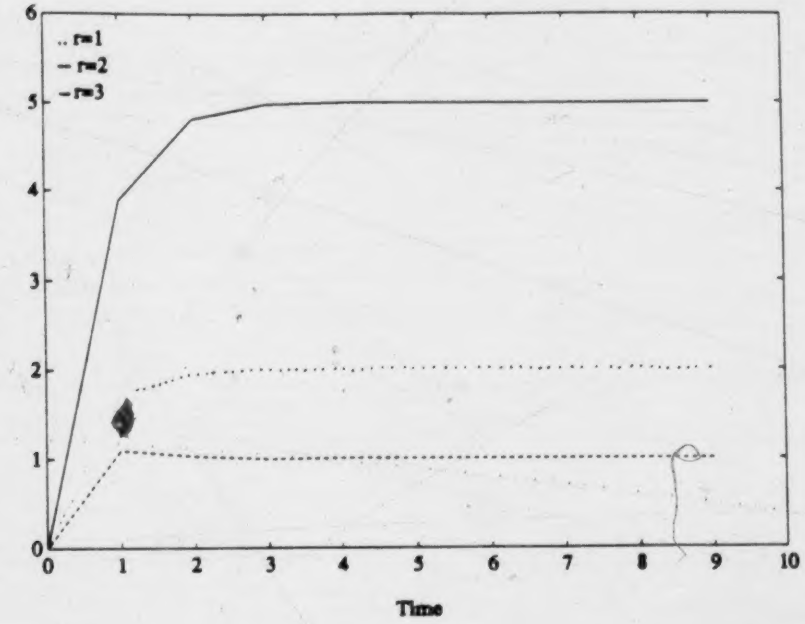

Figure 5.13a: Step Response for the Reduced Order Models Given in Example 5.4

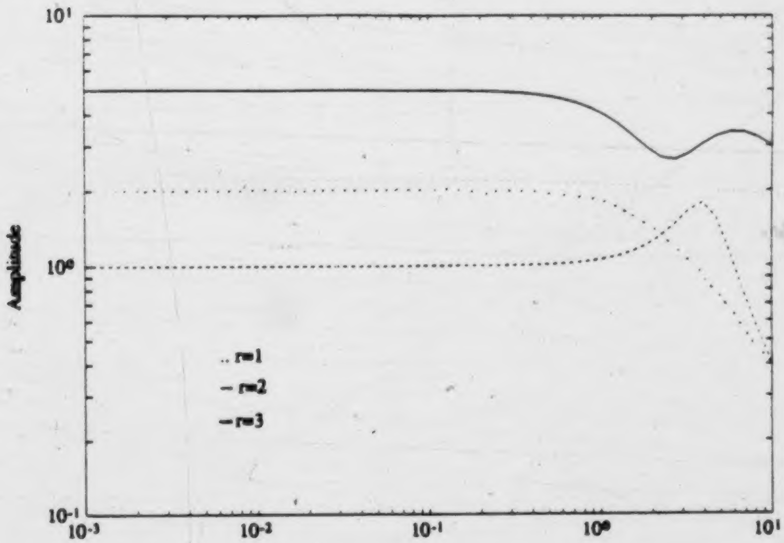

Figure 5.13b: Frequency Response for the Reduced Order Models Given in Example 5.4

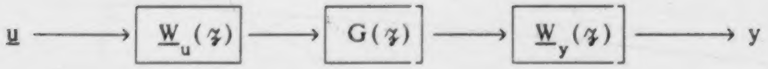

Figure 5.14: Frequency Weighted System

Grammian, $\underline{W}_c^2$, and the observability Grammian $\underline{W}_o^2$, will be greater than zero. However, for the third equation to have a solution, $\underline{X}$, the eigenvalues of $\underline{A}$ and $\underline{A}_w$ must be distinct (Al-Saggaf, 1987). The frequency weighted minimal realization can be transformed into an input-normalized, output-normalized or internally balanced frequency weighted realization via the appropriate linear transformation, $\underline{T}$. In the transformed realization $(\underline{\tilde{A}},\underline{\tilde{B}},\underline{\tilde{C}},\underline{\tilde{D}})$,

$$\underline{\tilde{A}} = \underline{T}^{-1}\underline{A}\underline{T}, \qquad [5.2.22a]$$

$$\underline{\tilde{B}} = \underline{T}^{-1}\underline{B}, \qquad [5.2.22b]$$

$$\underline{\tilde{C}} = \underline{C}\underline{T} \qquad [5.2.22c]$$

and

$$\underline{\tilde{X}} = \tilde{\underline{T}}^{-1}\underline{X}. \qquad [5.2.22d]$$

If the matrices are partitioned such that

$$\begin{bmatrix} \underline{\tilde{A}}_{11} & \underline{\tilde{A}}_{12} & \underline{\tilde{A}}_{13} \\ \underline{\tilde{A}}_{21} & \underline{\tilde{A}}_{22} & \underline{\tilde{A}}_{23} \\ \underline{\tilde{A}}_{31} & \underline{\tilde{A}}_{32} & \underline{\tilde{A}}_{33} \end{bmatrix} \quad \begin{bmatrix} \underline{\tilde{B}}_1 \\ \underline{\tilde{B}}_2 \\ \underline{\tilde{B}}_3 \end{bmatrix} \quad \begin{bmatrix} \underline{\tilde{C}}_1 & \underline{\tilde{C}}_2 & \underline{\tilde{C}}_3 \end{bmatrix} \qquad [5.2.23a]$$

$$\begin{bmatrix} \underline{\hat{X}}_1 \\ \underline{\hat{X}}_2 \\ \underline{\hat{X}}_3 \end{bmatrix} \qquad \qquad \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix} \qquad [5.2.23b]$$

where $\underline{\tilde{A}}_{11} \in \mathbb{R}^{rxr}$ and $\underline{\tilde{A}}_{22} \in \mathbb{R}^{kxk}$, Al-Saggaf (1987) developed a frequency weighted input normalized realization given by

$$\underline{G}_r(\mathfrak{z}) = \underline{D}_r + \underline{C}_r(\mathfrak{z}\underline{I}-\underline{A}_r)^{-1}\underline{B}_r \qquad [5.2.24a]$$

where

$$\underline{D}_r = [\underline{\check{C}}_2\underline{\hat{X}}_2 + \underline{\check{C}}_3\underline{\hat{X}}_3]\underline{C}_w \qquad [5.2.24b]$$

$$\underline{C}_r = \underline{\check{C}}_1 \qquad [5.2.24c]$$

$$\underline{A}_r = \underline{\tilde{A}}_{11} \qquad [5.2.24d]$$

$$\underline{B}_r = \underline{\hat{B}}_1 + [\underline{\tilde{A}}_{12}\underline{\hat{X}}_2+\underline{\tilde{A}}_{13}\underline{\hat{X}}_3]\underline{C}_w^{-1} \qquad [5.2.24e]$$

This frequency weighted realization will exhibit the following properties (Theorem 2, Al-Saggaf, 1987):

i) the reduced order model will be asymptotically stable

ii) If $\sigma_r(\underline{\Gamma}_{N,N}) > \sigma_{r+1}(\underline{\Gamma}_{N,N})$, the reduced order model will be observable

iii) All the poles of the frequency weighting are canceled by the zeros of the reduction error, $\underline{E}_{n-r}(\mathfrak{z}) = \underline{G}(\mathfrak{z})-\underline{G}_r(\mathfrak{z})$

iv) $\left\| \underline{E}_{n-r}(\mathfrak{z})\underline{W}(\mathfrak{z}) \right\|_\infty \leq 2^{*}\mathrm{tr}(\underline{\Sigma}_2)$ where the strict inequality holds if $\sigma_i \neq \sigma_{i+1}$ where $r \leq i \leq r+k+1$

5.2.3 Optimal Hankel Norm Model Reduction Techniques

The optimal model reduction problem is to find a process model of degree r<n for the process with degree n which will minimize a norm of the approximation error. Model reduction techniques which minimize the hankel norm of the error between the full order model (degree n) and a reduced order model (degree r),

$$\min \left\| \underline{G}(s) - \underline{G}_r(s) \right\|_H \qquad [5.2.26]$$

or

$$\min \left\| \underline{\Gamma} - \underline{\Gamma}_{r,r} \right\|_H \qquad [5.2.27]$$

have been developed. The hankel norm is defined as (Glover, 1984):

$$\left\| G(z) \right\|_H = \sigma_M(\underline{\Gamma}) = \sqrt{\lambda_{max}(\underline{W}_c^2 \underline{W}_o^2)} \qquad [5.2.28]$$

where $\underline{\Gamma}$ is the impulse hankel matrix for the process. The minimum hankel norm distance between $\underline{G}(z)$ and a r^{th} order approximation $\underline{G}_r(z)$ is bounded below by the r+1 second order mode of the system (Kung, 1981, Glover, 1984):

$$\left\| \underline{\Gamma}[\underline{G}(z)] - \underline{\Gamma}[\underline{G}_r(z)] \right\|_S \geq \sigma_{r+1}(\underline{\Gamma}) \qquad [5.2.29]$$

If an approximation matrix $\underline{G}_a(z)$ meets the hankel norm tolerance defined by

$$\left\| \underline{\Gamma}[\underline{G}(z)] - \underline{\Gamma}[\underline{G}_a(z)] \right\| \leq \rho \qquad [5.2.30]$$

where $\rho \in [\sigma_{r+1}, \sigma_r]$ then $\underline{G}_a(z)$ must have degree greater than or equal to r (Kung 1981).

In 1971, Adamjan, Arov and Krein (Silverman, 1980) investigated the hankel norm approximation problem for discrete scalar systems assuming an infinite impulse hankel matrix for the process is available. They showed that there will be more than one optimal r^{th} rank approximation of the infinite impulse hankel matrix of which at least one will be a hankel matrix. There will exist an unique hankel matrix, $\underline{\Gamma}_{r,r}$, of rank r which will approximate the infinite-dimensional system impulse hankel matrix, $\underline{\Gamma}$, such that the hankel norm $\left\|\underline{\Gamma} - \underline{\Gamma}_{r,r}\right\|_H$ is minimized over all bounded hankel matrices of rank r (Silverman, 1980). The optimal rth rank hankel matrix will be given by

$$\underline{\Gamma}_{r,r} = \underline{\Gamma} - \underline{\Gamma}(\phi_r(z)) \qquad [5.2.31a]$$

where

$$\phi_r(z) = \sigma_{r+1} \cdot \frac{\mu_{r+1}(z)}{\nu_{r+1}(z)} \qquad [5.2.31b]$$

and

$$\mu_i(z) = \sum_{j=1}^{\infty} \underline{z}_i(j) \cdot z^{-j} \qquad [5.2.31c]$$

$$\nu_i(z) = \sum_{j=1}^{\infty} \underline{v}_i(j) \cdot z^{j-1} \qquad [5.2.31d]$$

where $\underline{v}_i(j)$ and $\underline{z}_i(j)$ are the j^{th} components of the i^{th} right and left singular vectors, respectively, of the process impulse hankel matrix (Silverman, 1980). Adamjan also showed that $\nu_{r+1}(z)$ will have exactly r zeros inside the unit circle ($|z|<1$) when σ_{r+1} is a unique singular value of $\underline{\Gamma}$. In the frequency domain, there will exist a unique transfer function $G_r(z)$ which minimizes the hankel norm of the approximation error where

$$G_r(z) = G(z) - \phi_r(z) \qquad [5.2.32]$$

Furthermore, the error is given by

$$\left\|\underline{\Gamma} - \underline{\Gamma}_{r,r}\right\| = \left\|G(z) - G_r(z)\right\| = \sigma_{r+1}(\underline{\Gamma}) \qquad [5.2.33]$$

The restriction that the approximation matrix must be a hankel matrix does not affect the achievable error (Glover, 1984).

Silverman (1980) utilized balanced realizations to obtain a finite structure for the optimal hankel norm approximation developed by Adamjan in the form of a transfer function. Given a balanced realization of the nth order original system,

$$G(z) = \frac{n(z)}{d(z)} \qquad [5.2.34]$$

closed form expressions for $\nu_i(\mathfrak{z})$ and $\mu_i(\mathfrak{z})$ in equations 5.2.31c and 5.2.31d are developed. Since

$$\underline{\Gamma} = \underline{Z}\cdot\Sigma\cdot\underline{V}^T = \underline{Q}\cdot\underline{P} \qquad [5.2.35]$$

the rational functions in equations 5.2.31c and 5.2.31d can be rewritten as:

$$\mu_i(\mathfrak{z}) = \sigma_i^{-1/2}\underline{c}(\mathfrak{z}\underline{I} - \underline{A})^{-1}\underline{e}_i = \frac{\overset{*}{m}(\mathfrak{z})}{d(\mathfrak{z})} \qquad [5.2.36]$$

$$\nu_i(\mathfrak{z}) = \sigma_i^{-1/2}\underline{e}_i^T(\mathfrak{z}\underline{I} - \underline{A})^{-1}\underline{b} = \pm\frac{m^*(\mathfrak{z})}{d^*(\mathfrak{z})} = \pm\frac{\mathfrak{z}^{-(n-1)}m(\mathfrak{z}^{-1})}{\mathfrak{z}^{-n}\,d(\mathfrak{z}^{-1})} \qquad [5.2.37]$$

since $\underline{z}_i = \underline{Q}\cdot\Sigma^{-1/2}\underline{e}_i$ and $\underline{v}_i^T = \Sigma^{-1/2}\underline{P}\cdot\underline{e}_i$ (Silverman, 1980). It should be noted that $\mu(\mathfrak{z})/\nu(\mathfrak{z})$ is an all-pass function. An all-pass function is a function whose magnitude is unity at all frequencies. The r^{th} order transfer function which minimizes the hankel norm of the approximation error in equation 5.2.33 is given by the proper stable part of (Theorem 4.1, Silverman, 1980):

$$G_r(\mathfrak{z}) = \frac{p(\mathfrak{z})}{m^*(\mathfrak{z})} \qquad [5.2.38a]$$

where

$$p(\mathfrak{z}) = \frac{n(\mathfrak{z})m^*(\mathfrak{z}) - \lambda d^*(\mathfrak{z})m(\mathfrak{z})}{d(\mathfrak{z})} \qquad [5.2.39]$$

and $\lambda=\pm\sigma_{r+1}$. $G_r^*(\mathfrak{z})$ will not be an optimal approximation to $G(\mathfrak{z})$ in the L^∞ norm sense unless $r = n-1$ (Silverman, 1980). If $r \neq n-1$, then

$$\left\| G(\mathfrak{z}) - G_r^*(\mathfrak{z}) \right\|_\infty \neq \sigma_{r+1} \qquad [5.2.39]$$

Bettayeb (1980) used a similar approach to Silverman (1980) to develop an algorithm for obtaining optimal approximations of any specified order for continuous time, SISO system when the original system is asymptotically stable and internally balanced. There will exist a unique hankel operator of rank r, $\underline{\Gamma}_r$, which minimizes the spectral norm of the error $\left\|\underline{\Gamma} - \underline{\Gamma}_r\right\|_S$ over all bounded hankel operators of rank r such that (Bettayeb, 1980):

$$\left\|\underline{\Gamma} - \underline{\Gamma}_r\right\|_S = \sigma_{r+1}(\underline{\Gamma}) \qquad [5.2.40]$$

Furthermore,

$$\underline{\Gamma}_r = \underline{\Gamma} - \sigma_{r+1}\underline{\Gamma}\{\phi_{r+1}(s)\} \qquad [5.2.41a]$$

where

$$\phi_{r+1}(s) = \frac{\underline{z}_{r+1}(s)}{\underline{v}_{r+1}(s)} = \frac{\underline{c}(s\underline{I}-\underline{A})^{-1}e_{r+1}}{e^T_{r+1}(-s\underline{I}-\underline{A})^{-1}\underline{b}} \qquad [5.2.41b]$$

and e_i is the i^{th} unit vector (Bettayeb, 1980). If $(\underline{A},\underline{b},\underline{c})$ is a balanced realization of G(s), then $|\phi_{r+1}(s)|=1$ for all $s=j\omega$. Therefore, the optimal r^{th} order approximation is given by

$$G_r(s)=\underline{c}(s\underline{I}-\underline{A})^{-1}\underline{b}-\phi_{r+1}(s)\cdot\frac{\underline{c}(s\underline{I}-\underline{A})^{-1}e_{r+1}}{e^T_{r+1}(-s\underline{I}-\underline{A})^{-1}\underline{b}} \qquad [5.2.42]$$

$G_r(s)$ is not necessarily stable and therefore may not be optimal. If $r \neq n-1$, the L^∞ norm of the approximation error may not be minimal. However, if $\sigma_{r+1} \ll \sigma_r$, then

$$\left\|G(s) - G_r(s)\right\|_\infty \cong \sigma_{r+1}(\underline{\Gamma}). \qquad [5.2.43]$$

A discrete version of the approximation algorithm has also been derived (Silverman, 1980).

Kung (1981) utilized the approximation theory of Adamjan (1971) to develop a solution to the minimum degree approximation (MDA) problem for scalar discrete systems. The minimum degree approximation problem involves the determination of an approximation impulse hankel matrix, $\underline{\Gamma}_{app}$, such that the minimum degree bound given in equation 5.2.30 is achieved with the preassigned tolerance $\rho \in [\sigma_{r+1},\sigma_r]$. In this case, the rank of $\underline{\Gamma}_{app}$ will be greater than or equal to r (Kung, 1980).

Kung (1981) extended their scalar MDA algorithm approach to the multivariable case for discrete time systems using block impulse hankel matrices. Kung assumes that Γ has a singular value of multiplicity p and that the tolerance, ρ, is exactly that singular value ($\rho=\sigma_{k+1}=\sigma_{k+2}=...=\sigma_{k+p}$, and $\sigma_k \geq \rho \geq \sigma_{k+p+1}$). With this assumption, there will be p independent pairs of singular vectors corresponding to the singular value ρ. A k^{th} order approximation is obtained from these p pairs of singular vectors. The MDA (minimal degree approximation) theorem involves three steps:

1. computation of a minimal basis solution of the polynomial matrix equation
2. solution of an algebraic Ricatti equation
3. partial fraction expansion of the rational matrix

The solution to the MDA problem is quite time consuming. The first step in the algorithm is solved using a Gauss elimination type procedure or a fast projection method. The second step requires an iterative procedure to solve. The third step is partial fraction expansion and is the most time intensive portion of the algorithm. Therefore, Kung and Lin's method will be difficult to implement and may not be practical.

Glover (1984) developed a continuous time algorithm for obtaining optimal hankel norm approximations for multivariable causal transfer functions $\underline{G}(s)$ of degree n by $\underline{G}_r(s)$ of degree r. Glover derived a complete characterization of all approximations $\underline{G}_r(s)$ which minimize the hankel norm of the approximation error $\underline{G}(s) - \underline{G}_r(s)_H$ using balanced realizations and all-pass transfer functions. It is shown that all solutions are simple functions of a balanced realization. An all-pass transfer function is one for which

$$G(s)G^{+}(-\bar{s})=I \qquad [5.2.44]$$

and all the singular values of the impulse hankel matrix (second order modes) are unity. Glover (1984) also derives a bound on the L^{∞} norm of the approximation error for one class of optimal hankel-norm approximations.

Although Glover's algorithm is derived for a continuous system, the method can be utilized for discrete systems. The bilinear transformation

$$z = \frac{1+s}{1-s} \qquad [5.2.45]$$

can be used to transform the discrete system to a continuous form. Once the optimal hankel norm approximation is found, the solution is converted back to the discrete form using the bilinear transformations (Glover, 1984, Wahlberg, 1986)

$$s = \frac{z-1}{z+1} \qquad [5.2.46]$$

The conversion from continuous time to discrete time to continuous time can be done because the singular values of the system impulse hankel matrix (the second order modes) are invariant under linear transformation (Wahlberg, 1986).

Glover (1984) derives a class of solutions to the optimal hankel norm approximation problem. The hankel norm is used as an error criteria to approximate a causal transfer function matrix, $\underline{G}(s)$, by an optimal anticausal transfer function matrix of degree r (all poles are unstable and in the right half plane). If $\underline{G}(s)$ is a rational (mxm) transfer function matrix, then $\underline{G}_r(s)$, of degree ≤r, will be an optimal hankel norm approximation to $\underline{G}(s)$ if and only if there exists an anticausal function $\underline{F}(s)$ (degree ≤n+r-1) such that (Theorem 7.2, Glover 1984):

$$\underline{E}(s) = \underline{G}(s) - \underline{F}(s) = \underline{G}(s) - \underline{G}_r(s) - \underline{G}_{us}(s) \qquad [5.2.47a]$$

satisfies

$$\underline{E}(s)\underline{E}^T(-\bar{s}) = \sigma^2_{r+1}I \qquad [5.2.47b]$$

where $\underline{E}(s)$ is the error matrix. Therefore, the error matrix, $\underline{E}(s)$, will be all-pass. In this case, the hankel norm of the approximation error will be minimal,

$$\left\|\underline{G}(s)-\underline{G}_r(s)\right\|_H = \sigma_{r+1} \qquad [5.2.48]$$

Let $(\underline{A},\underline{B},\underline{C},\underline{D})$ be a balanced realization of $\underline{G}(s)$ where $\underline{A} \in \mathbb{R}^{n\mathrm{x}n}$, $\underline{B} \in \mathbb{R}^{n\mathrm{x}m}$, $\underline{C} \in \mathbb{R}^{p\mathrm{x}n}$ and $\underline{D} \in \mathbb{R}^{m\mathrm{x}m}$ with second order modes

$$\sigma_1 \geq \sigma_2 \geq \ldots \sigma_r \geq \sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_{r+k} \geq \sigma_{r+k+1} \geq \ldots \geq \sigma_n > 0 \qquad [5.2.49]$$

The controllability and observability Grammians, $\underline{W}_c^2$ and $\underline{W}_o^2$, can be partitioned as

$$\underline{W}_c^2 = \begin{bmatrix} \underline{\Sigma}_1 & \underline{0} \\ \underline{0} & \sigma \underline{I}_{r+1} \end{bmatrix} = \underline{W}_o^2 \qquad [5.2.50a]$$

where

$$\underline{\Sigma}_1 = \mathrm{diag}(\ \sigma_1,\ \sigma_2, \ldots,\ \sigma_r,\ \sigma_{r+2},\ \sigma_{r+3}, \ldots,\ \sigma_n) \qquad [5.2.50b]$$

$\sigma \neq 0$ and

$$\delta(\underline{\Sigma}_1^2 - \sigma^2 \underline{I}_r) = 0 \qquad [5.2.50c]$$

$\delta(\cdot)$ represents the number of eigenvalues on the imaginary axis. Similarly, the state matrices can be partitioned such that

$$\underline{A} = \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} \\ \underline{A}_{21} & \underline{A}_{22} \end{bmatrix} \quad \underline{B} = \begin{bmatrix} \underline{B}_1 \\ \underline{B}_2 \end{bmatrix} \quad \underline{C} = \begin{bmatrix} \underline{C}_1 \\ \underline{C}_2 \end{bmatrix} \qquad [5.2.52]$$

An optimal $\underline{G}_r(s)$ is given by the stable part of

$$\underline{F}(s) = \underline{G}_r(s) + \underline{G}_{us}(s) = \hat{\underline{D}} + \hat{\underline{C}}(s\underline{I}-\hat{\underline{A}})^{-1}\hat{\underline{B}} \qquad [5.2.53]$$

where $\hat{\underline{D}}$, $\hat{\underline{C}}$, $\hat{\underline{A}}$, and $\hat{\underline{B}}$ are given by (Theorem 6.3, GLover, 1984):

$$\hat{\underline{A}} = \eta^{-1}(\sigma_{r+1}^2 \underline{A}_{11}^T + \underline{\Sigma}_1 \underline{A}_{11} \underline{\Sigma}_1 - \sigma_{r+1} \underline{C}_1^T \underline{U}\ \underline{B}_1^T) \qquad [5.2.54a]$$

$$\hat{\underline{B}} = \eta^{-1}(\underline{\Sigma}_2 \underline{B}_1 + \sigma_{r+1} \underline{C}^T \underline{U}) \qquad [5.2.54b]$$

$$\hat{\underline{C}} = \underline{C}_1 \underline{\Sigma}_1 + \sigma_{r+1} \underline{U} \underline{B}_1^T \qquad [5.2.54c]$$

$$\hat{\underline{D}} = \underline{D} - \sigma_{r+1} \underline{U} \qquad [5.2.54d]$$

where $\underline{U}$ is a unitary matrix which satisfies

$$\underline{B}_2 = -\underline{C}_2^T \underline{U} \qquad [5.2.54e]$$

and

$$\eta = \underline{\Sigma}_1^2 - \sigma_{r+1}^2 \underline{I} \qquad [5.2.54f]$$

Glover (Theorem 6.1, 1984) showed that if one wished to approximate a causal transfer function, $\underline{G}(s)$, by an anticausal function, $\underline{E}(s)$, the smallest L^{∞} error that can be achieved is the hankel norm of $\underline{G}(s)$:

$$\|\underline{G}(s)\|_{H} \leq \|\underline{G}(j\omega) - \underline{E}(j\omega)\|_{L^{\infty}} \qquad [5.2.46]$$

The class of solutions described above can be generalized to include non-square systems by relaxing the restriction that $\underline{U}$ must be unitary ($\underline{U}^{T}\underline{U} \leq \underline{I}$). Also, $\underline{U} = -\underline{C}_{2}\underline{B}_{2}^{-1}$ (Corollary 7.3, Glover, 1984).

Glover (1984) further extends the results to find all optimal hankel norm approximations $\underline{E}(s) = \underline{G}_{r}(s)+\underline{G}_{us}(s)$ to $\underline{G}(s)$ such that

$$\frac{\|\underline{G}(s) - \underline{G}_{r}(s) - \underline{G}_{us}(s)\|}{\|\underline{G}(s)\|_{H}} \qquad [5.2.49]$$

is an all-pass function. The $\underline{D}$ matrix is not determined by the optimal hankel norm approximation method since the hankel norm is independent of $\underline{D}$ (Glover, 1984). However, Glover (1984) uses an L^{∞} norm error bound to choose a $\underline{D}$. Glover showed that there will exist a constant $\tilde{\underline{D}}$ such that (Corollary 9.3, Glover, 1984):

$$\|\underline{G}(j\omega)+\underline{D}\|_{L^{\infty}} \leq (\sigma_{1}+\sigma_{2}+\ldots+\sigma_{n}) \qquad [5.2.50]$$

and

$$\|\underline{G}(j\omega)\|_{L^{\infty}} \leq 2(\sigma_{1}+\sigma_{2}+\ldots+\sigma_{n}) \qquad [5.2.51]$$

Li (1987) and Jonckheere (1987) looked at hankel norm reduction techniques to approximate spectra. They developed a phase approximation algorithm using a balanced realization of the phase of the spectral factor E(s) to construct a reduced order spectral factor $\hat{E}(s)$ using the method of Glover (1984) such that

$$\left\| \frac{w}{w^{*}} - \frac{w}{\hat{w}^{*}} \right\| \qquad [5.2.52]$$

is small in the hankel norm sense where w(·) is a random process spectral factor. They also derived L^{∞} error bounds on the approximation such that

$$\left\| E(s)-\hat{E}(s) \right\|_{\infty} = \left\| \frac{w}{w^{*}} - \frac{w}{\hat{w}^{*}} \right\|_{\infty} \leq 4(\sigma_{r+1} + \dots + \sigma_{N}) \qquad [5.2.53]$$

and the hankel norm error bound on the approximation is

$$\left\| E(s) - \hat{E}(s) \right\|_{H} \leq 2(\sigma_{r+1} + \dots + \sigma_{N}) \qquad [5.2.54]$$

σ_i are the hankel singular values of the stable projection of E(s).

Other researchers have extended the results for the optimal hankel norm problem to investigate the H^{∞} optimization problem

$$\inf_{K(s) \in RH\infty} \left\| [R(s)+K(s)]M(s,\gamma_o)^{-1} \right\|_{\infty} = 1 \qquad [5.2.55]$$

Chang (1987) presents a review of the H^{∞} optimization techniques. Chang recommends Glover's method for solving the hankel norm optimization problem which will construct an optimal K(s).

5.3 Summary

Several techniques which utilize the impulse hankel matrix to identify a state space model of the process have been evaluated. Although the controllability and observability Grammians are equal, they are not diagonal matrices. Therefore, the realizations are balanced so they will exhibit good conditioning with respect to controllability and observability. However, they are not optimal.

All four identification methods produce an accurate process model of a deterministic system if enough impulse response parameters are used in the algorithms and the correct model order is known. As the number of impulse response parameters used in the identification decreases from N, where N is the settling time/sampling time, the models identified using Kung's algorithm exhibit an increasing steady state offset. The frequency range over which the models are valid decreases as the high and low frequency modes are successively neglected. The models identified with the other three algorithms will identify a full order model.

If N impulse response parameters are used in the identification of a reduced order model, all four algorithms will give realizations which exhibit a steady state offset. As the model order decreases, the identified models neglect the high frequency modes of the system. As the number of impulse response parameters used in the identification decreases, the low frequency modes are neglected and the steady state offset increases.

If there is noise in the system, the realizations identified from the Ho-Kalman algorithms will exhibit a steady state offset and neglect the high frequency modes of the system. As the noise level decreases, the steady state offset will decrease. The modified algorithm of Hajdasinski produces the "best" model.

Several techniques which reduce the order of a balanced realization were evaluated. Truncation of balanced realizations will produce good reduced order models if the singular values of the components of the state space which are eliminated are relatively small. However, if the components which are neglected contribute significantly to the state space, the reduced order models will not adequately describe the process. Frequency weighting can be used to minimize the model error within certain frequency ranges by choosing an appropriate input and/or output filter. The method of Glover produces a model which minimizes the hankel norm of the error.

The identification and order reduction techniques based on SVD are numerically well conditioned. Even in the presence of noise, the systems will remain well-conditioned if balanced realization are used to describe the process.

Chapter 6 SVD in the Analysis and Design of Feedback Control Strategies for Multivariable Systems

6.0 Introduction

The design procedure for feedback control strategies involves formulation or specification of the control objectives, selection of the best control configuration and design of an appropriate controller. Practical applications of feedback strategies in chemical processes which have multiple inputs and outputs are complicated due to the presence of process interactions. Closed loop interactions result because a single output can be influenced by more than one input and/or a single input can affect more than one process output. These interactions can result in tuning and stability problems for control schemes. In chemical plant regulatory control applications, the primary control objective is the maintenance of the controlled variables (outputs) at their setpoints in the presence of disturbances. Therefore, a practical, regulatory control scheme will minimize the effects of disturbances and process interactions in the system while maintaining a stable and robust closed loop control strategy. For servo control strategies, the main objective of the control system is to closely track a changing setpoint with the process output. Therefore, process interactions should be minimized such that a change in an input variable will significantly affect only one output variable.

There are several methods described in the literature which have been used to analyze process interactions in multivariable systems: Rijinsdorp interaction quotient (Rijinsdorp, 1965), relative gain array (Bristol, 1966, McAvoy, 1983), dynamic relative gain array (Bristol, 1977, Witcher, 1977), relative dynamic gain array (Tung, 1977, 1981), characteristic loci (MacFarlane, 1977), average dynamic gain array (Gagnepain, 1982), direct Nyquist array (Jensen, 1981, 1986), inverse Nyquist array (Rosenbrock, 1969), direct gain interaction matrix (Johnston, 1984), and several others. These analysis techniques do not require knowledge of the controller structure. There are other methods available such as the interaction index (Davison, 1969, 1970) and the IMC interaction measure Economou's (1986) which apply to specified controller structures. However, Jensen (1981) showed that the IMC measure is identical to the normalized direct Nyquist array. Jensen (1981) presents a good review of most of the currently available methods. Lieslehto and Koivo (1987) developed an expert system to analyze process interactions. The expert system selects an appropriate analysis procedure based on a given process model (transfer function vs state space) and provides an indication of the degree of interaction in the system. If the interactions are small, then the expert system will select the "best" input-output pairings for multiloop control. In this chapter, it is shown how singular value decomposition of the system gain matrix can be used to analyze interactions between manipulated variables and controlled variables. A method of "measuring" these interactions was also developed. However, it is shown that the measure is similiar to the direct Nyquist array and the IMC interaction measure.

A common feedback control structure used in industry for multivariable systems is a multiloop structure. In this design, the loop "pairings" of the manipulated variables and the controlled variables are chosen such that the open loop process interactions are minimized. Each input-output pair defines a distinct control loop. It should be noted that a single control loop could involve a single output and a linear combination of several inputs. For systems with significant interactions, decouplers or compensators, based on a model of the process, can be used to compensate for interactions. Once the process has been adequately decoupled, diagonal controllers, usually PID type, are designed using the "optimum" variable pairings to minimize the closed loop interactions. Conventional SISO tuning procedures can be utilized for each "control loop".

McAvoy (1983), Moore (1986), Lau (1985), Bequette (1986), Johnston (1984) and several others have utilized singular value analysis techniques to design multiloop control schemes. Using the components from the SVD of the system transfer function matrices, analysis can be carried out to:

1. determine the "best" input-output pairings to minimize the open-loop interactions
2. obtain a relative measure of open-loop and closed loop process interactions
3. obtain a relative measure of the controllability of the system
4. determine the optimum sensor locations and choice of manipulated variables in large systems to minimize interactions and maximize the sensitivity of the control loops. The sensitivity of the control loop refers to the gain between a given input and a corresponding output.

Control system design and analysis methods which use singular value analysis possess the "good" numerical properties associated with singular value decomposition. In addition, these methods can be used in a straightforward manner for dynamic analysis of both square and non-square systems. However, SVD techniques are dependent on the scaling of the process variables.

Scaling can change the singular values and vectors of the transfer function matrices and can change the condition number. This is the main weakness of these methods. In the following sections, these methods are described and evaluated using simulated examples.

Several investigators have adopted multivariable control strategies which "decouple" or compensate for the process interactions within the controller itself. Control techniques such as Dynamic Matrix Control (DMC) (Cutler, 1977, 1982), Quadratic DMC (Garcia, 1986), Generalized Predictive Control (GPC) (Mohtadi, 1986), and Internal Model Control (IMC) (Garcia, 1982, 1985a, 1985b, Rivera, 1986) avoid the explicit input-output pairing requirements of the multiloop control schemes. These multivariable methods are long-range predictive control strategies which require "good" process models to ensure stability and robustness. As the dimensions of the multivariable systems increase, the complexity of these multivariable controllers also increases.

6.1 Multiloop Control System Design

6.1.1 Process Interactions

The design of a stable and robust multiloop control scheme requires an understanding of the nature of process interactions in the open loop and closed loop systems. This information can be used to decide on the best control configurations (input-output pairings), assist in the design of suitable decouplers (compensators), and give an indication of the potential difficulties in terms of controllability of the proposed control scheme.

Jensen (1981, 1986) analyzed the signal transmissions between the input and output variables in closed loop multivariable systems in terms of the sum of direct, parallel, interactive and disturbance transmissions. For a general (nxn) feedback system, shown in Figure 6.1, these transmittances are defined as:

1. Direct ($r_i \rightarrow y_i$)
2. Parallel ($r_i \rightarrow y_i$ via other loops)
3. Interactive ($r_j \rightarrow y_i$,$j \neq i$)
4. Disturbance ($\xi_k \rightarrow y_i$)

where r_i is the i^{th} reference input variable, y_i is the i^{th} controlled variable, and ξ_k is the k^{th} disturbance variable as shown in Figure 6.2. The parallel transmittances occur because a portion of the input signal, r_i, can enter the other control loops, travel around these loops and return to the original control loop via the interaction paths.

For the system in Figure 6.1,

$$\underline{y} = (\underline{I} + \underline{G}_p\underline{K}\underline{G}_c\underline{H})^{-1}\underline{G}_p\underline{K}\underline{G}_c \cdot \underline{r} + (\underline{I} + \underline{G}_p\underline{K}\underline{G}_c\underline{H})^{-1}\underline{G}_L \cdot \underline{\xi} \qquad [6.1.1]$$

The Laplace transform variable, s, has been omitted in equation 6.1.1 and all subsequent equations in this section for simplicity. The "effective" process TFM, $\underline{Q}$, is given by:

$$\underline{Q} = \underline{G}_p\underline{K} \qquad [6.1.2]$$

where $\underline{K}$ is a precompensator or decoupler. The return difference matrix for the system is given by:

$$\underline{F} = \underline{I} + \underline{Q}\underline{G}_C \qquad [6.1.3]$$

Substitution of equations 6.1.2 and 6.1.3 into equation 6.1.1 when $\underline{H} = \underline{I}$ gives:

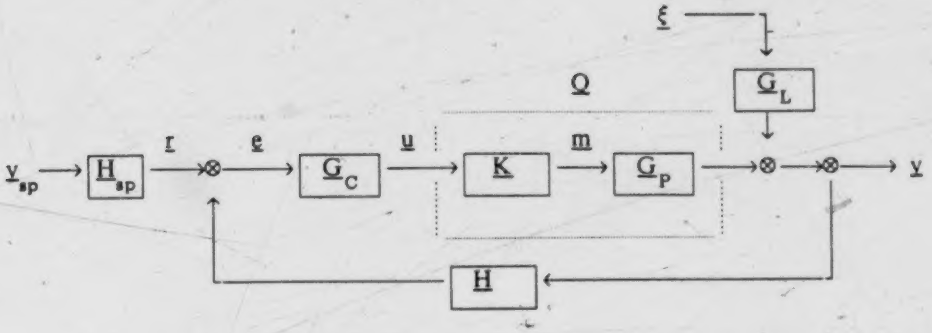
ξ
Q
G_L
y_sp
H_sp
r
e
G_C
u
K
m
G_P
y
H

Figure 6.1: Closed Loop Process Diagram Where $\underline{\xi}$ = Disturbance Vector (qx1), $\underline{u}$ = Control Input Vector (nx1), $\underline{m}$ = Manipulated Variable Vector, $\underline{e}$ = Feedback Error Vector $\underline{r}$ = Reference Input Vector, vector, $\underline{y}_{sp}$ = Setpoint Vector (nx1), $\underline{y}$ = Output Vector (nx1)

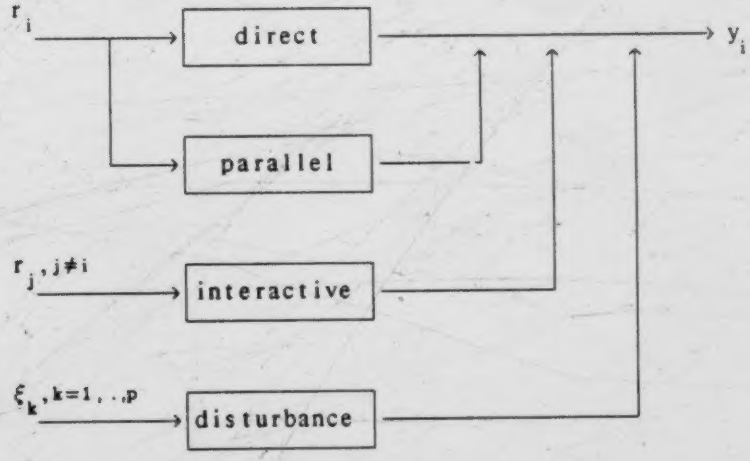
r_i
direct
parallel
r_j, j≠i
interactive
ξ_k, k=1, . ., p
disturbance
y_i

Figure 6.2: Closed Loop Signal Transmittances

$$\underline{y} = \underline{E}^{-1}\underline{Q}\underline{G}_{C}\cdot\underline{r} + \underline{E}^{-1}\underline{G}_{L}\cdot\underline{\xi} = \underline{G}^{*}\cdot\underline{r} + \underline{G}_{L}^{*}\cdot\xi \qquad [6.1.4]$$

where

$$\underline{E}^{-1} = \frac{\mathrm{adj}(\underline{E})}{\det(\underline{E})} = \frac{\{c_{ij}^{T}\}}{\det(\underline{E})} \qquad [6.1.5]$$

and $c_{ij} = (-1)^{i+j}\underline{M}_{ij}$ are the cofactors of $\underline{E}$ and $\underline{M}_{ij}$ are the corresponding minors of $\underline{E}$. Each element in the closed loop transfer function matrix, $\underline{G}^{*}$, can be expressed as,

$$g_{ij}^{*} = \sum_{l=1}^{n} \frac{c_{li}k_{j}q_{lj}}{\det(\underline{E})} \qquad [6.1.6]$$

where n is the number of control loops (the number of outputs). A similar expression can be derived for the elements in the closed loop disturbance transfer function matrix $\underline{G}_{L}^{*}$. Therefore, the i^{th} controlled variable is determined from (Jensen, 1981):

$$y_i = g_{ii}^{*}\cdot r_i + \sum_{\substack{j=1\\ j\neq i}}^{n} g_{ij}^{*}\cdot r_j + \sum_{k=1}^{p} g_{Lij}^{*}\cdot\xi_k$$

$$= \underbrace{\frac{k_i q_{ii}}{\det(\underline{E})} r_i}_{\text{direct}} + \underbrace{\sum_{l=1}^{n}\frac{(c_{li}-\delta_{li})k_i q_{li}}{\det(\underline{E})} r_i}_{\text{parallel}} + \underbrace{\sum_{\substack{j=1\\ j\neq i}}^{n}\sum_{l=1}^{n}\frac{c_{li}k_j q_{lj}}{\det(\underline{E})} r_j}_{\text{interactive}} + \underbrace{\sum_{j=1}^{p}\left[\sum_{l=1}^{n}\frac{c_{li}g_{lj}^{L}}{\det(\underline{E})}\right]\xi_j}_{\text{disturbance}} \qquad [6.1.7]$$

where δ_{li} is the unit impulse function corresponding to the i^{th} output in the l^{th} control loop.

Jensen (1981) showed that the closed loop transmission paths, given in equation 6.1.7, could be estimated using open loop information. If

$$\underline{P}(s) = \underline{Q}(s)\cdot\underline{G}_{C}(s) \qquad [6.1.8]$$

represents the open loop TFM, then the elements in the closed loop TFM have a common denominator and the numerator can be expressed in terms of the sum of the minors of order 1 to (n-1) of the open loop TFM as

$$\dot{g}_{ij} = \frac{1}{\det(\underline{E})} \cdot \left[P^{i}_{j} + \sum_{l_1=1}^{n} P^{il_1}_{jl_1} + \sum_{l_1=1}^{n} \sum_{l_2=l_1+1}^{n} P^{il_1l_2}_{jl_1l_2} + \cdots \right.$$

$$\left. + \sum_{l_1=1}^{n} \sum_{l_2=l_1+1}^{n} \cdots \sum_{l_n=l_{n-1}+1}^{n} P^{il_1l_2\ldots l_n}_{jl_1l_2\ldots l_n} + \cdots + \delta_{ij}\det(P) \right] \qquad [6.1.9]$$

where $P^{il_1l_2..l_n}_{jl_1l_2..l_n}$ is the minor of $\underline{P}$ obtained by deleting all rows except those indicated by the superscripts and deleting all columns except those indicated by the subscripts (Jensen, 1981). The k^{th} order transmittances can be expressed in terms of the k^{th} order minors of $\underline{P}$.

As can be seen from equation 6.1.7, the relative magnitudes of the transmittances can be compared by examining the numerators of the expressions because all of the terms are divided by a common factor, det($\underline{E}$). The first order minor representing the direct transmittances between r_i and y_i is given by the diagonal element of the open loop TFM, $k_i q_{ii}$. The first order interactive transmittance between r_j and y_i, j≠i, is given by the off diagonal element of the open loop TFM, $k_j q_{ij}$. Therefore, the relative magnitude of the closed loop first order interaction transmittances with respect to the direct transmittance can be approximated by comparing the off diagonal elements of the open loop TFM to the corresponding diagonal elements (Jensen, 1981). Generally, the open loop TFM is structured such that the diagonal elements are larger than the off-diagonal elements. Therefore, as the number of off-diagonal elements through which a given input signal passes increases, the transmittances should decrease. Jensen (1981) concluded that for most practical systems, the magnitude of the minors of the open loop TFM and the magnitude of the transmittances will generally decrease as their order increases. Therefore, the first order interaction transmittance, represented by the first order minor of the open loop TFM, should provide a reasonable approximation to the total closed loop interaction. For the 2×2 system shown in Figure 6.3, the elements of the closed loop transfer function matrix are given by,

$$\dot{g}_{11} = \frac{k_1 q_{11} - k_2 q_{12} k_1 q_{21} + k_1 q_{11} k_2 q_{22}}{\det(\underline{E})} \qquad [6.1.10a]$$

$$\dot{g}_{12} = \frac{k_2 q_{12}}{\det(\underline{E})} \qquad [6.1.10b]$$

$$\dot{g}_{21} = \frac{k_1 q_{21}}{\det(\underline{E})} \qquad [6.1.10c]$$

$$\dot{g}_{22} = \frac{k_2 q_{22} - k_1 q_{21} k_2 q_{12} + k_1 q_{11} k_2 q_{22}}{\det(\underline{E})} \qquad [6.1.10d]$$

The actual transmission paths through the closed loop system are given by the minors of the open loop transfer function matrix divided by the determinant of $\underline{E}$. The division by a polynomial results in an infinite number of transmission paths. In a feedback system, a portion of the input signal, r_i, can enter the j^{th} control loop ($j=1,..n,\ j \neq i$), pass N (1 to ∞) times around these loops and return to the i^{th} control loop generating an infinite number of parallel transmittance paths. In a similar manner, there will be an infinite number of interactive and disturbance transmittances paths due to r_j ($j=1,..n, j \neq i$) and ξ_k($k=1,..,p$), respectively. For the 2×2 system in Figure 6.3, the physical closed loop transmission paths are given by the infinite series,

$$\dot{g}_{11} = k_1q_{11} - k_1q_{21}k_2q_{12} + k_1q_{21}k_2q_{12}k_1q_{21}k_2q_{12} - \cdots \quad [6.1.11a]$$

$$\dot{g}_{12} = k_2q_{12} - k_2q_{12}k_1q_{11} + k_2q_{12}k_1q_{21}k_2q_{12} - \cdots \quad [6.1.11b]$$

$$\dot{g}_{21} = k_1q_{21} - k_1q_{21}k_2q_{22} + k_1q_{21}k_2q_{12}k_1q_{21} - \cdots \quad [6.1.11c]$$

$$\dot{g}_{22} = k_2q_{22} - k_2q_{12}k_1q_{21} + k_2q_{12}k_1q_{21}k_2q_{12}k_1q_{21} - \cdots \quad [6.1.11d]$$

The closed loop parallel, interactive and disturbance transmissions can be influenced by more than one off-diagonal element in the closed loop TFM and therefore, more than one element in the open loop TFM. By definition, the direct transmittances will be influenced by the diagonal elements of the closed loop TFM only and thus, only the diagonal elements of the open loop TFM. For the 2×2 system in Figure 6.3, the primary transmission paths can be identified. The primary paths are ones in which a signal traverses any section of the path only once (Jensen, 1981). The primary direct transmittance between r_1 and y_1 is k_1q_{11} (1st order minor). The primary interactive transmittance between r_1 and y_2 is k_1q_{21} (1st order minor), and the primary interactive transmittance between r_2 and y_1 is k_2q_{12} (1st order minor). The primary parallel transmittance between r_1 and y_1 (r_2 and y_2) is $-k_1q_{21}k_2q_{12}$.

In multivariable systems, the design objectives are to maximize the sum of the direct plus parallel transmittances but minimize the interactive and disturbance transmittances. However, these two objectives are not independent because the primary parallel transmission is the product of the primary interactive transmittances. From equation 6.1.1 and 6.1.7, it can

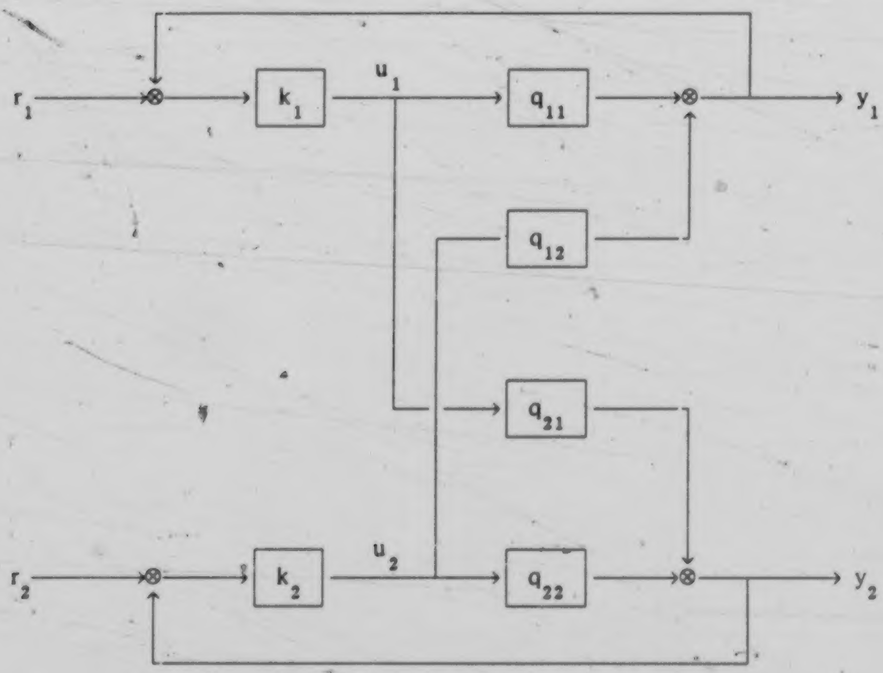
r_1
k_1
u_1
q_{11}
y_1
q_{12}
q_{21}
r_2
k_2
u_2
q_{22}
y_2

Figure 6.3: 2x2 Closed Loop System with Multiloop Control Scheme. k_1 and k_2 are the Feedback Controller Gains

be seen that the interactive and disturbance transmittances in the closed loop system will approach zero as the feedback gains, k_i, approach infinity. The direct and parallel transmittances will approach unity with high feedback gains.

6.1.2 Singular Value Analysis of Process Interactions

Relative measures of open loop and closed loop interactions, as defined by Jensen (1981), and the optimum input-output pairings for a multiloop strategy can be obtained from singular value analysis of the open loop transfer function matrix. The linear time-invariant MIMO system, shown in Figure 6.2, (neglecting disturbances) is described by,

$$\underline{y}(s) = \underline{G}_p(s)\underline{m}(s) = \underline{Q}(s)\underline{u}(s) \qquad [6.1.12]$$

where $\underline{y}(s) \in \mathbb{R}^p$, $\underline{u}(s) \in \mathbb{R}^m$, and $\underline{G}_p, \underline{Q} \in \mathbb{R}^{p\times m}$. Applying singular value decomposition to the process TFM, $\underline{Q}(s)$, in equation 6.1.12 gives

$$\underline{y}(s) = \underline{Z}(s)\underline{\Sigma}(s)\underline{V}^T(s)\underline{u}(s) = \sum_{i=1}^{p} \sigma_i(s)\cdot\underline{z}_i(s)\cdot\underline{v}_i^T(s)\underline{u}(s) \qquad [6.1.13]$$

The columns of $\underline{V} = [\underline{v}_1\ \underline{v}_2\ \dots\ \underline{v}_m]$ and $\underline{Z} = [\underline{z}_1\ \underline{z}_2\ \dots\ \underline{z}_p]$ form orthonormal bases which span the input and output spaces, respectively.

If the basis vectors, $\underline{v}_i$ and $\underline{z}_i$, for the input and output spaces, respectively, are aligned with the standard Euclidean basis vectors, then an input entering in the direction of $\underline{e}_i$ will be scaled by σ_i and reappear at the output in the direction of $\underline{e}_i$. This corresponds to a non-interacting loop where the signal y_i is completely determined by u_i and u_i affects only y_i. If

$$\underline{V}(s) = \underline{Z}(s) = \underline{E} = \underline{I} \qquad [6.1.14]$$

then

$$\underline{y}(s) = \underline{Q}(s)\underline{u}(s) = \underline{\Sigma}(s)\underline{u}(s) \qquad [6.1.15]$$

Since $\underline{\Sigma}(s)$ is diagonal, the plant is naturally decoupled. If $\underline{v}_i \neq \underline{e}_i$ and $\underline{z}_i \neq \underline{e}_i$, the system will exhibit some degree of interaction depending on the alignment of $\underline{z}_i$ and $\underline{v}_i$ with $\underline{e}_i$.

Singular value decomposition provides a method of identifying open loop and closed loop signal transmittances. Each element in the process TFM can be written as the sum of dyads,

$$q_{ij}(s) = \sum_{l=1}^{m} \sigma_l(s)[\underline{e}_j^T \underline{z}_l(s)][\underline{v}_l^+(s)\underline{e}_i] \qquad [6.1.17a]$$

$$q_{ij}(s) = \sum_{l=1}^{n} \sigma_l(s) z_{il}(s) v_{jl}^+(s) \qquad [6.1.17b]$$

The vector products $(\underline{e}_j^T \underline{z}_l)$ and $(\underline{v}_l^+ \underline{e}_i)$ represent the projection of the l^{th} left singular vector onto the $\underline{e}_j$ standard basis vector, z_{jl}, and the projection of the l^{th} right singular vector onto the $\underline{e}_i$ standard basis vector, v_{il}^+, respectively. The elements of the closed loop TFM in terms of the singular vectors and values of the process TFM are given by,

$$g_{ij}^{*} = \sum_{l=1}^{p} \sum_{k=1}^{p} \frac{c_{li} k_j \sigma_k z_{ik} v_{jk}^+}{\det(\underline{E})} \qquad [6.1.18]$$

For the i^{th} control loop, the open loop signal transmittances are given by

$$y_i = k_i q_{ii} e_i + \sum_{\substack{j=1 \\ j \neq i}}^{m} k_j q_{ij} e_j = \sum_{l=1}^{p} k_i \sigma_l z_{il} v_{il}^+ e_i + \sum_{l=1}^{p} \sum_{\substack{j=1 \\ j \neq i}}^{m} k_j \sigma_l z_{il} v_{jl}^+ e_j$$

$$= \underbrace{k_i \sigma_i z_{ii} v_{ii}^+ e_i}_{\text{direct}} + \underbrace{\sum_{\substack{l=1 \\ l \neq i}}^{p} k_i \sigma_l z_{il} v_{il}^+ e_i}_{\text{parallel}} + \underbrace{\sum_{l=1}^{p} \sum_{\substack{j=1 \\ j \neq i}}^{m} k_j \sigma_l z_{il} v_{jl}^+ e_j}_{\text{interactive}} \qquad [6.1.19]$$

A 2x2 closed loop system, incorporating the singular value decomposition components of the process TFM, is shown in Figure 6.4. The area contained within the dotted area represents the process Q. Primary direct transmittances between r_1 (u_1) and y_1 pass through v_{11}, are scaled by a "gain", σ_1, and go through z_{11}. Primary parallel transmittances between r_1 (u_1) and y_1 go through v_{12}, are scaled by "gain σ_2, and then pass through z_{12} before exiting at y_1. Primary interactive transmittances due to u_2 pass through v_{22}, are scaled by "gain" σ_2, go through z_{12}, then exit through y_1. In addition, a portion of the r_2 (u_2) signal travels through v_{21}, is scaled by"gain σ_1, passes through z_{11} and appears at the output y_1. Similarly, for output y_2, the primary direct, parallel and interactive transmittances can beshown to be $k_2\sigma_2 z_{22} v_{22}$, $k_2\sigma_1 z_{21} v_{21}$, and $k_1(\sigma_1 z_{21} v_{11} + \sigma_2 z_{22} v_{12})$, respectively. Therefore, the open loop equation can be written as,

$$\begin{aligned}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \begin{bmatrix} \sigma_1 z_{11} v^{+}_{11} + \sigma_2 z_{12} v^{+}_{12} & \sigma_1 z_{11} v^{+}_{21} + \sigma_2 z_{12} v^{+}_{22} \\ \sigma_1 z_{21} v^{+}_{11} + \sigma_2 z_{22} v^{+}_{12} & \sigma_1 z_{21} v^{+}_{21} + \sigma_2 z_{22} v^{+}_{22} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &= \begin{bmatrix} k_1\sigma_1 z_{11} v^{+}_{11} + k_1\sigma_2 z_{12} v^{+}_{12} & k_2\sigma_1 z_{11} v^{+}_{21} + k_2\sigma_2 z_{12} v^{+}_{22} \\ k_1\sigma_1 z_{21} v^{+}_{11} + k_1\sigma_2 z_{22} v^{+}_{12} & k_2\sigma_1 z_{21} v^{+}_{21} + k_2\sigma_2 z_{22} v^{+}_{22} \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}\end{aligned} \qquad [6.1.20]$$

Jensen's interaction analysis shows that the open loop interactions or the primary interactive transmittance paths, given in equation 6.1.19, provide a reasonable indication of first order closed loop interactions. Therefore, interaction measures based on singular value analysis of the process TFM which compare the relative magnitude of the open loop direct and parallel transmittances,

$$k_i \sigma_i z_{ii} v^{+}_{ii} e_i + \sum_{\substack{l=1 \\ l \neq i}}^{n} \sigma_i z_{il} v^{+}_{il} e_i \qquad [6.1.21]$$

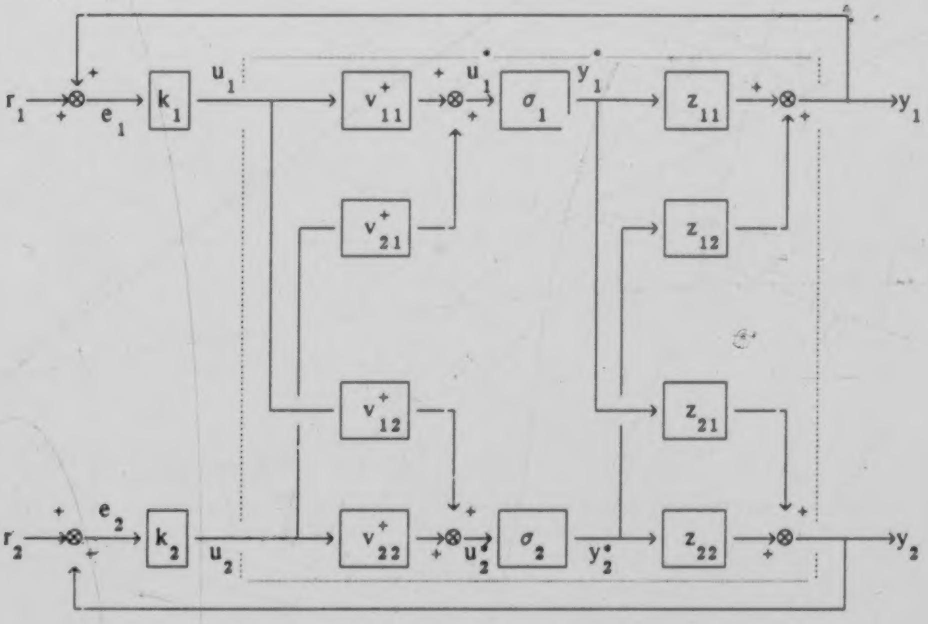
r_1
e_1
k_1
u_1
v^+_{11}
u^*_1
σ_1
y^*_1
z_{11}
y_1
v^+_{21}
z_{12}
v^+_{12}
z_{21}
r_2
e_2
k_2
u_2
v^+_{22}
u^*_2
σ_2
y^*_2
z_{22}
y_2

Figure 6.4: A 2x2 Multiloop Control System Showing the SVD of the Process TFM

to the magnitude of the open loop interaction transmittance

$$\sum_{l=1}^{p} \sum_{\substack{j=1 \\ j\neq i}}^{m} k_j \sigma_l z_{il} v^{+}_{jl} e_j \qquad [6.1.22]$$

should be "good" measures of open loop and approximate first order closed loop interaction for loop pairing $u_i \Rightarrow y_i$ in the i^{th} control loop. Interaction measures based on SVA will be discussed further in section 6.1.4.5.

6.1.3 Optimum Controlled and Manipulated Variable Pairings for Multiloop Control Schemes Using SVA

The multiloop control system design problem involves the determination of the best controlled and manipulated variable pairings for each control loop. The "optimum" variable pairings for a particular process should minimize the interactions in the uncompensated plant and produce satisfactory closed loop performance over the frequency band of characteristic disturbances. The manipulated and controlled variable pairings, $u_i \rightarrow y_i$, for each control loop should be chosen such that the primary direct signal transmittance components, v_{ii} and z_{ii}, are larger in magnitude than the interactive components, v_{jk} and z_{ik} ($j\neq i$, $1\leq k\leq p$) and the parallel components, v_{ik} and z_{ik} ($k\neq i$, $1\leq k\leq p$) (Jensen, 1981). Therefore, the element corresponding to the largest absolute value in each column of $\underline{Z}$, $\underline{z}_i$, should be paired with the element corresponding to the largest absolute value in each column of $\underline{V}$, $\underline{v}_i$, (Moore, 1986). These variable pairings should minimize the open loop interactions for the uncompensated plant since the most sensitive manipulated variable is paired with the most sensitive controlled variable associated with the loop gain σ_i.

The first column of $\underline{Z}$ ($\underline{z}_1$), corresponding to the largest decoupled process "gain" (σ_1), indicates the combination of controlled variables which are most affected by the system. Therefore, $\underline{z}_1$ is the easiest direction in which the system can be changed. Similarly, the first column of $\underline{V}$ ($\underline{v}_1$) indicates the combination of control actions which will have the greatest effect on the process. The last columns in $\underline{Z}$ ($\underline{z}_n$) and $\underline{V}$ ($\underline{v}_m$), corresponding to the smallest decoupled process "gain", indicate the combination of controlled variables least affected by the system and the combination of control actions which have the least effect on the system, respectively.

Lau (1985) used a similar method to obtain the "best" input-output pairings. The open loop equation, 6.1.13, can be written in terms of the sum of rotational matrices, $\underline{W}$, as,

$$\underline{y}(s) = \sum_{l=1}^{n} \sigma_l(s) \cdot \underline{z}_l(s) \cdot \underline{v}_l^T(s) \cdot \underline{u}(s)$$

$$= \sum_{l=1}^{n} \sigma_l(s) \cdot \underline{W}_l(s) \cdot \underline{u}(s) \qquad [6.1.23]$$

Each loop (or node) in the system has a scaling factor $\sigma_l(s)$ and a rotational transformation, $\underline{W}_l(s)$, between the input and the output, associated with it. The maximum entry in the rotational matrix (max $|w_{ij}^l|$) defines a $u_i \rightarrow y_j$ pairing for the lth loop. Since $\underline{W}_l = \underline{z}_l \underline{v}_l^+$, the largest element in $\underline{W}_l$ will correspond to the largest elements in $\underline{z}_l$ and $\underline{v}_l$.

Example 6.1: The open loop transfer function matrix for a system of two interacting tanks, assuming linear valves on all flow lines, is derived in Appendix A. The control objective in this system is to control the liquid levels, h_1 and h_2, in the tanks by manipulating the inlet flows to the two tanks, q_1 and q_2. If the tank cross-sectional areas and the valve resistances are:

$$A_1 = A_2 = 1 \text{ m}^2 \tag{6.1.24}$$

$$R_1 = R_2 = 1 \text{ s/m}^2 \tag{6.1.25}$$

the process TFM is,

$$\underline{G}_p(s) = \begin{bmatrix} \dfrac{s+2}{s^2+3s+1} & \dfrac{1}{s^2+3s+1} \\ \dfrac{1}{s^2+3s+1} & \dfrac{s+1}{s^2+3s+1} \end{bmatrix} \tag{6.1.26}$$

The SVD components of the steady state gain matrix, $\underline{G}_p(0)$ are,

$$\underline{Z} = \begin{bmatrix} 0.851 & -0.526 \\ 0.526 & 0.851 \end{bmatrix} \qquad \underline{V} = \begin{bmatrix} 0.851 & -0.526 \\ 0.526 & 0.851 \end{bmatrix}$$

$$\underline{\Sigma} = \begin{bmatrix} 2.618 & 0 \\ 0 & 0.382 \end{bmatrix} \tag{6.1.27}$$

The $\underline{Z}$ and $\underline{V}$ matrices are equal in this example because the TFM is symmetric. Matching the largest elements in the columns of $\underline{Z}$ with the largest elements in the columns of $\underline{V}$ gives the loop pairings:

$$h_1 \rightarrow q_1 \tag{6.1.28a}$$

$$h_2 \rightarrow q_2 \tag{6.1.28b}$$

with decoupled loop gains of 2.618 and 0.382, respectively. These loop pairings are intuitively obvious if the physical arrangement of the system is examined. A sketch of the system is given in appendix A.

Example 6.2: A binary distillation column at the University of Alberta was modeled using the MOD modeling program (Langman, 1987). The top composition (x_D) and the bottom composition (x_B) are to be controlled by manipulating the reflux flow rate (R) and the steam flow rate to the reboiler (S). The distillation column is highly non-linear with varying time delays in the TFM. Langman (1987) derived an "average" model which approximated the non-linear behaviour:

$$\begin{bmatrix} x_T \\ x_B \end{bmatrix} = \begin{bmatrix} \dfrac{1.42e^{-s}}{7.86s+1} & \dfrac{-0.669e^{-3.8s}}{15.0s+1} \\ \dfrac{2.29e^{-6.5s}}{19.9s+1} & \dfrac{-4.54e^{-5.3s}}{18.5s+1} \end{bmatrix} \cdot \begin{bmatrix} R \\ S \end{bmatrix}$$

where

$$\underline{G}_p(0) = \begin{bmatrix} 1.420 & -0.669 \\ 2.290 & -4.540 \end{bmatrix} \qquad [6.1.30]$$

SVD of $\underline{G}_p(0)$ yields:

$$\underline{Z} = \begin{bmatrix} 0.224 & -0.970 \\ 0.970 & 0.224 \end{bmatrix} \qquad \underline{V} = \begin{bmatrix} 0.490 & -0.872 \\ -0.872 & 0.490 \end{bmatrix}$$

$$\underline{\Sigma} = \begin{bmatrix} 5.238 & 0 \\ 0 & 0.938 \end{bmatrix} \qquad [6.1.31]$$

Matching the absolute value of the largest components of the columns of $\underline{Z}$ with the largest components of the columns of $\underline{V}$ gives the optimum loop pairings,

$$x_B \Rightarrow S \qquad [6.1.32]$$

$$x_D \Rightarrow R \qquad [6.1.33]$$

with decoupled loop gains of 5.238 and 0.938, respectively. This analysis implies that the bottom composition from the distillation column will be more sensitive to the steam flow to the reboiler than the sensitivity of the top composition to reflux flow rate.

A dynamic analysis of the optimum variable pairings can be performed over a frequency range using SVA. At each frequency, the process TFM is evaluated and the SVD of the matrix is computed. The dynamic analysis will indicate whether the optimum loop pairings are frequency dependent. If the optimum pairings remain the same over a large low frequency range and only change at high frequencies, a decoupler designed at a low frequency may be adequate for stable control. However, if the optimum pairings change over the frequency range, a dynamic decoupler will be required to ensure good performance over the entire bandwidth.

Example 6.3.: Tung (1977) considered a 2×2 numerical example where the static and dynamic optimum loop pairings are different. The process TFM is given by,

$$\underline{G}_p(s) = \frac{\begin{bmatrix} (s+6.45)(s+1.55) & -5(s+3) \\ 2(s+3) & (s+3-2i)(s+3+2i) \end{bmatrix}}{(s+4+2i)(s+4-2i)(s+3)} \qquad [6.1.34]$$

At steady state, SVD of $\underline{G}_p(0)$ gives,

$$\underline{Z} = \begin{bmatrix} 0.952 & 0.307 \\ -0.307 & 0.952 \end{bmatrix} \qquad \underline{V} = \begin{bmatrix} 0.411 & 0.912 \\ -0.912 & 0.411 \end{bmatrix}$$

$$\underline{\Sigma} = \begin{bmatrix} 0.312 & 0 \\ 0 & 0.161 \end{bmatrix} \qquad [6.1.35]$$

Matching the elements with the largest absolute value in the columns of $\underline{Z}$ with the elements corresponding to the largest absolute value in the columns of $\underline{V}$ results in loop pairings:

$$u_2 \rightarrow y_1 \qquad [6.1.36a]$$

$$u_1 \rightarrow y_2 \qquad [6.1.36b]$$

These loop pairings would be "optimum" at steady state. Substituting $s=i\omega$ into equation 6.1.34 gives

$$\underline{G}_p(i\omega) = \frac{\begin{bmatrix} (i\omega + 6.45)(i\omega + 1.55) & -5(i\omega+3) \\ 2(i\omega+3) & (i\omega+3-2i)(i\omega+3+2i) \end{bmatrix}}{(i\omega+4+2i)(i\omega+4-2i)(i\omega+3)} \qquad [6.1.37]$$

At ω =10 Hz, the SVD components are,

$$\underline{Z} = \begin{bmatrix} 0.952 & 0.307 \\ -0.307 & 0.952 \end{bmatrix} \qquad \underline{V} = \begin{bmatrix} 0.411 & 0.912 \\ -0.912 & 0.411 \end{bmatrix}$$

$$\underline{\Sigma} = \begin{bmatrix} 0.312 & 0 \\ 0 & 0.161 \end{bmatrix} \qquad [6.1.38]$$

In this case, the optimum variable pairings are

$$u_1 \Rightarrow y_1 \qquad [6.1.39a]$$

$$u_2 \Rightarrow y_2 \qquad [6.1.39b]$$

which are different than the steady state pairings. Figure 6.5 shows the variable pairings obtained by matching the element with the largest absolute value in each column of $\underline{Z}(i\omega)$ with the element corresponding to the largest absolute value in each column of $\underline{V}(i\omega)$ over a frequency range. At frequencies less than 6 Hz, the optimum pairings are $u_2 \Rightarrow y_1$ and $u_1 \Rightarrow y_2$ for the first and second control loops, respectively. However, at frequencies greater than 7 Hz, the optimum loop pairings are $u_1 \Rightarrow y_1$ and $u_2 \Rightarrow y_2$. At ω = 6 Hz, the loop pairings are reversed. Therefore, a compensator designed at low frequencies will not take into account the directionality of the system at high frequencies. As a consequence, the control system performance will deteriorate at high frequencies.

There are other methods available for determining the optimum variable pairings for multiloop or decoupled control schemes. The most commonly used tool is the relative gain array (RGA) developed by Bristol (1966) which is scaling independent. For a linear, time-invariant system with an equal

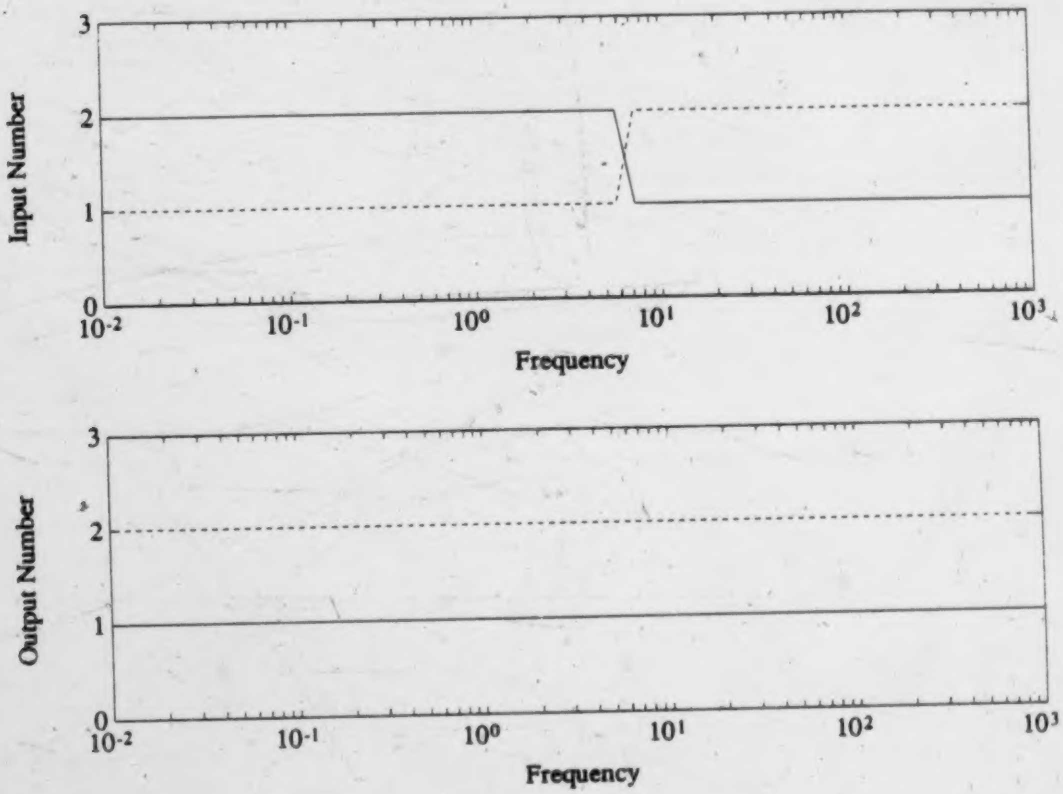

Figure 6.5: Optimum Input/Output Pairings Versus Frequency for Example 6.3 (Tung, 1977)

number of controlled and manipulated variables, optimum variable pairings can be determined directly from the RGA. Since the RGA involves ratio quantities, it is invariant to scaling of the process variables. The SVA techniques discussed previously are scaling dependent but can be applied to non-square systems unlike the RGA. The RGA method is a steady state method which is easy to use since it requires only steady state gains. However, if there are strong interactions existing during transient periods of operation, the RGA analysis could give "bad" variable pairings because the high frequency dynamics have been neglected. Witcher (1977), Bristol (1977), Tung (1977, 1981), and Gagnepain (1982) extended the RGA to include the effects of process dynamics. However, the dynamic methods are more complicated to calculate than SVA. Johnston (1984) developed the direct gain matrix which determines dynamic optimum variable pairings. This method can be used for square and non-square systems to select the optimum variable pairings. In addition, DGM can be used to reduce the number of manipulated variables in ill-conditioned systems and to improve the performance and sensitivity of systems through the use of internal control loops. However, the DGM is scaling dependent.

6.1.4 Process Interaction Measures

When the optimum variable pairings for the uncompensated plant have been determined, an interaction measure can be used to indicate the relative magnitudes of interactions and determine which loops require decoupling to obtain satisfactory control performance. There are several interaction measures which have been published in the literature. Jensen (1981) presents a very good summary of the various interaction measures and their relationship to closed loop interactions.

6.1.4.1 Relative Gain Array (RGA)

Bristol (1966, 1977) developed the relative gain array as a tool to assist in the design of multiloop control strategies. The elements in the RGA, μ_{ij}, are defined as the ratio of the open loop gain between u_j and y_i when all the other loops are closed and under perfect control. If μ_{ij} = 1.0, the system is said to be completely decoupled. If μ_{ij} = 0.5, the system is said to be highly interacting. If $\mu_{ij} > 1$, all the elements in the relative gain matrix will be of comparable magnitude indicating significant interactions in the system. The manipulated and controlled variables corresponding to positive RGA elements closest to one are paired to form a feedback loop to minimize interactions. The method is easy to use and requires only steady state process gains. However, Jensen (1981) showed that the RGA does not provide a direct measure of closed loop interactions. The RGA elements are a steady state measure of the limiting values of the ratio of direct transmittances to the sum of direct and parallel transmittances for a system with one loop open and the others closed.

6.1.4.2 Direct Nyquist Array (DNA)

The direct Nyquist array (DNA) (Jensen, 1981) provides a direct measure of open loop process interactions and an approximate measure of first order closed loop process interactions. Each element in the direct Nyquist array (DNA) is a polar plot representing the magnitude of the corresponding element in the open loop TFM evaluated over a specified frequency range. Gershgorin circles, evaluated at each frequency, can be superimposed onto the diagonal elements of the DNA to give a measure of interaction. The radii of the Gershgorin circles are equal to the sum of the magnitudes of all the off diagonal elements in the i^{th} column of the process transfer matrix

$$(r_i(\omega) = \sum_{j=1}^{n} |q_{ji}(i\omega)| \ , i \neq j) \qquad [6.1.40]$$

The center of the circles are located on the $g_{ii}(i\omega)$ locus. A comparison of the magnitude of the diagonal elements to the corresponding radii of the Gershgorin circles at each frequency is a measure of the effect of a process input u_i on the output y_i as compared to the effect on y_j ($j \neq i$). The rows of the process transfer function matrix can also be used to compute Gershgorin circles. The interaction measure in this case will be a measure of the relative influence that an input u_i has on a particular output y_i as compared to the influence that u_j ($j \neq i$) has on y_i.

To minimize closed loop interactions, the magnitudes of the diagonal elements of the closed loop TFM should be larger than the off-diagonal elements. From Jensen's (1981) analysis on process interactions, the requirement for minimal closed loop interactions implies that the largest elements in the open-loop transfer function matrix should also be on the diagonal. The optimum input-output pairings are chosen such that the largest DNA displays are on the diagonal of the array and the Gershgorin circles are as small as possible. The interaction measure from the DNA can be used directly to design a control scheme using Nyquist exact loci techniques or characteristic loci techniques. The performance and robustness of the control system and the control quality can be analyzed using related frequency domain techniques. The only disadvantage of this method is that it is scaling dependent.

6.1.4.3 IMC Interaction Measure

Economou (1986) introduced two interaction measures derived from an internal model control (IMC) design approach. The row IMC measure is defined as:

$$R_i(i\omega) = \sum_{\substack{j=1 \\ j\neq i}}^{n} |g_{ij}(i\omega)| \Big/ \sum_{i=1}^{n} |g_{ij}(i\omega)| \qquad [6.1.41]$$

and characterizes the relative importance of an input assigned to the loop (on an output y). The column IMC measure is defined as:

$$C_i(i\omega) = \sum_{\substack{j=1 \\ j\neq i}}^{n} |g_{ji}(i\omega)| \Big/ \sum_{j=1}^{n} |g_{ji}(i\omega)| \qquad [6\ 1.42]$$

and characterizes the interactions imposed by the i^{th} loop on the other loops. This measure is essentially a normalized version of the Gershgorin circles where the radii and the magnitude of $g_{ii}(s)$ has been normalized by dividing by $\sum|g_{ij}|$ for R_i and similarly by $\sum|g_{ji}|$ for C_i (Jensen, 1986).

Therefore, this measure is a normalized DNA method presented in the form of a Bode plot (Jensen, 1981).

6.1.4.4 Characteristic Loci Measure

MacFarlane (1977) used the misalignment of vector spaces to define a measure of interaction for use in their characteristic loci (CL) design method for multivariable systems. In the CL design, the set of frequency dependent eigenvectors (characteristic vectors) and eigenvalues (characteristic gains) of the open loop TFM, $\underline{P}(s)$ given in equation 6.1.8, are manipulated such that a desired closed loop response is obtained. The characteristic loci are defined as the loci traced out by the eigenvalues (characteristic gains) of $\underline{P}(s)$ in the complex plane as s traverses the standard Nyquist contour. The controller matrix, $\underline{G}_c(j\omega)$, is chosen such that the set of characteristic loci satisfies the generalized Nyquist stability criterion.

$\underline{P}(s)$ can be expressed in terms of the dyads formed from the characteristic gains and the characteristic directions as

$$\underline{P}(s) = \underline{W}(s)\Lambda(s)\underline{X}^T(s) = \sum_{i=1}^{n} \lambda_i(s)\underline{w}_i(s)\underline{x}_i^+(s) \qquad [6.1.43]$$

where λ_i is the i^{th} characteristic gain, $\underline{w}_i$ is the i^{th} characteristic direction and $\underline{x}_i$ forms the dual basis to the characteristic directions such that

$$\underline{X} = (\underline{W}^{-1})^T \qquad [6.1.44]$$

The closed loop TFM becomes (MacFarlane, 1977):

$$\underline{R}(s) = [\underline{I}_n + \underline{P}(s)]^{-1}\underline{P}(s) = \sum_{i=1}^{n}\left[\frac{\lambda_i(s)}{1+\lambda_i(s)}\right]\cdot w_i(s)x_i^T(s) \qquad [6.1.45]$$

Therefore, a reference input, $\underline{r}(s)$, entering the closed loop system will be split into components directed along the characteristic vector set $\{\underline{w}_i(s)$: $i=1,..,n)$. The i^{th} component will be scaled by the closed loop gain of

$$\frac{\lambda_i(s)}{1+\lambda_i(s)} \qquad [6.1.46]$$

The components of the output space will be directed along the same set of characteristic direction vectors as the components in the input space.

If all the characteristic gains are large over the bandwidth of the system, $\lambda_i(j\omega) >> 1$, the closed loop gains will approach unity and

$$\begin{aligned}\underline{y}(j\omega) &= \underline{R}(j\omega)\cdot\underline{r}(j\omega)\\ &\cong \underline{W}(j\omega)\underline{X}^T(j\omega)\cdot\underline{r}(j\omega)\\ \underline{y}(j\omega) &\simeq \underline{r}(j\omega)\end{aligned} \qquad [6.1.47]$$

since $\underline{W}\cdot\underline{V}^T = \underline{I}$. The system will exhibit little closed loop interaction and therefore good setpoint tracking. At low frequencies, $\underline{G}_c(j\omega)$ can be chosen such that the gains of the characteristic loci, $\lambda_i(j\omega)$, have sufficiently high gain over the operating bandwidth to provide the required accuracy for tracking of $\underline{r}(j\omega)$ by $\underline{y}(j\omega)$ (MacFarlane, 1977). In this case, the interactions in the closed loop system will be small. At high frequencies, the Nyquist stability criterion usually requires that the characteristic gains be small (MacFarlane, 1977). Therefore, as the frequency approaches infinity, the characteristic loci of $\underline{P}(j\omega)$ will usually approach zero such that

$$\underline{I}_n + \underline{P}(j\omega) \rightarrow \underline{I}_n \qquad [6.1.48]$$

Therefore,

$$\underline{R}(j\omega) \rightarrow \underline{P}(j\omega) \qquad [6.1.49]$$

The interactions in the closed loop system at high frequencies are essentially the same as in the open loop system. To eliminate high frequency interactions, the open loop interactions must be modified (MacFarlane, 1977).

If the characteristic vectors set of $\underline{R}(s)$, $\{\underline{w}_i(s): 1 \leq i \leq p\}$ and therefore of $\underline{P}(s)$ is closely aligned with the standard Euclidean basis set, $\{\underline{e}_i: 1 \leq i \leq p\}$, the closed loop system will exhibit minimal interactions since $\underline{W} = \underline{X} = \underline{I}$ and

$$\begin{aligned} \underline{y}(s) &= \underline{R}(s) \cdot \underline{r}(s) \\ &= (\underline{I} + \underline{W}(s)\Lambda(s)\underline{X}^T(s))^{-1}\underline{W}(s)\Lambda(s)\underline{X}^T(s) \cdot \underline{r}(s) \end{aligned}$$

$$\underline{y}(s) = (\underline{I} + \Lambda(s))^{-1}\Lambda(s) \cdot \underline{r}(s) = \underline{D}(s) \cdot \underline{r}(s) \qquad [6.1.50]$$

Since $\underline{D}(s)$ is a diagonal matrix, an input $r_i(s)$ will affect only $y_i(s)$. As the vectors $\underline{w}_i(s)$ and $\underline{e}_i$ become misaligned, the magnitude of the off-diagonal elements in $\underline{W}(s)$ increase from zero increasing the effect of $r_j(s)$ ($j \neq i$) on $y_i(s)$. The magnitude of the diagonal elements in $\underline{W}(s)$ will decrease from one decreasing the effect of $r_i(s)$ on $y_i(s)$. Therefore, the open loop interactions and the high frequency closed loop interactions will increase. The angle (MacFarlane, 1977)

$$\phi_i(j\omega) = \cos^{-1}\left\{ \frac{|(\underline{w}_i(j\omega), \underline{e}_i)|}{\underline{w}_i(j\omega)} \right\} \qquad [6.1.51]$$

between the vectors $\underline{w}_i(s)$ and $\underline{e}_i$ can be used to indicate the misalignment of these vectors. Therefore, plots of the characteristic gains of $\underline{P}(j\omega)$ and the alignment angles, $\phi_i(j\omega)$, as a function of frequency can be used to analyze process interactions and the performance of the closed loop system over the operating bandwidth of the system (MacFarlane, 1977). The optimum loop pairings for the i^{th} loop are chosen such that $\underline{w}_i(s)$ is the eigenvector which makes the smallest angle with the standard basis vector $\underline{e}_i$. The eigenvalue of $\underline{P}(s)$ corresponding to $\underline{w}_i(s)$ will be a reasonable approximation of the transference of the closed loop signals from $r_i(j\omega)$ to $y_i(j\omega)$ at a particular frequency (MacFarlane, 1977).

If the eigenvectors are approximately orthogonal (not excessively skewed), then high characteristic gains, $\lambda_i(s)$, and small alignment angles will ensure good closed loop performance over the system operating bandwidth. However, if the characteristic vector set is excessively skewed (not orthogonal), a complete set of eigenvectors which span the output space does not exist. Therefore, the system gains and directions can not be completely defined. In addition, the condition number of the system will be large and the analysis will be very sensitive to errors in the process model. In these systems, the eigenvectors and eigenvalues can not be used to assess closed loop performance. Alternatively, principal gains (singular values) and principal (singular) vectors of the open loop TFM can be used to determine the system gains and directions (MacFarlane, 1977). If the principal gains of $\underline{P}(s)$ are sufficiently high, $\underline{y}(j\omega)$ will closely track $\underline{r}(j\omega)$ regardless of the directions of the eigenvectors of $\underline{P}(s)$. If the singular vectors or principal vectors are closely aligned with the standard Euclidean basis vectors, the closed loop interactions will be minimal.

6.1.4.5 SVA Interaction Measures

Lau (1985a) defines a dynamic measure of interaction which represents the alignment angle between the singular vector dyads and the Euclidean standard basis dyads of the input and output spaces. This interaction measure indicates the degree of alignment of the l^{th} loop rotational matrix $\underline{W}_l$, given in equation 6.1.23, with respect to an ideal natural loop rotational matrix, $\overline{\underline{W}}_l$. An ideal natural loop is defined as a single loop which interacts minimally with other loops in the system. The basis vectors of the input and output spaces for an ideal loop are aligned with the standard basis vectors. For the l^{th} loop, with loop pairing $u_i \rightarrow y_i$, the rotational matrix $\underline{W}_l$ is given by,

$$\underline{z} = \begin{bmatrix} z_{1i} \\ z_{2i} \\ \vdots \\ z_{pi} \end{bmatrix} \quad \underline{v} = \begin{bmatrix} v_{1i} \\ v_{2i} \\ \vdots \\ v_{mi} \end{bmatrix} \quad \underline{W} = \begin{bmatrix} z_{1i}v^{+}_{1i} & z_{1i}v^{+}_{2i} & \cdots & z_{1i}v^{+}_{mi} \\ z_{2i}v_{1i} & z_{2i}v_{2i} & \cdots & z_{2i}v_{mi} \\ \vdots & \vdots & & \vdots \\ z_{pi}v^{+}_{1i} & z_{pi}v^{+}_{2i} & \cdots & z_{pi}v^{+}_{mi} \end{bmatrix} \qquad [6.1.52]$$

The corresponding ideal loop rotational matrix, $\overline{\underline{W}}_l$, with loop pairing $u_i \rightarrow y_i$, is the projection of $\underline{W}_l$ onto the standard Euclidean dyad, $e_{ii} = \underline{e}_i \underline{e}_i^T$, and is given by (Lau, 1985a)

$$\overline{\underline{z}}_i = \begin{bmatrix} 0 \\ \vdots \\ z_{ii} \\ \vdots \\ 0 \end{bmatrix} \quad \overline{\underline{v}}_i = \begin{bmatrix} 0 \\ \vdots \\ v_{ii} \\ \vdots \\ 0 \end{bmatrix} \quad \overline{\underline{W}}_l = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & z_{ii}v^{+}_{ii} & \cdots \\ 0 & 0 \cdots & 0 \cdots & 0 \end{bmatrix} \qquad [6.1.53]$$

Lau's interaction measure for the l^{th} loop, Θ, is defined as,

$$\cos \Theta_l = \frac{\overline{\underline{W}}_l}{\underline{W}_l} = \sqrt{w_{ii} \cdot w_{ii}^T} = |(\underline{e}_i^T \underline{z}_i)(\underline{v}_i^+ \underline{e}_i)| = |z_{ii} v_{ii}^+| \qquad [6.1.54]$$

since $\underline{W}_l$ =1.0. Lau also defines a total interaction measure where each nodal interaction measure is weighted by its corresponding singular value. The angle Θ indicates the deviation of $\underline{P}(j\omega)$ from an ideal decoupled system,

$$\cos \Theta = \left[\frac{\sum_{i=1}^{n} \sigma_i^2 (\cos \Theta_i)^2}{\sum_{i=1}^{n} \sigma_i^2} \right]^{1/2} \qquad [6.1.55]$$

Lau's interaction measure compares the relative magnitude of the (i,i) component of $\underline{W}_l$ with the other elements in $\underline{W}_l$. However, the elements in $\underline{W}_l$ represent only the signal transmittances which enter the system and are scaled by σ_i corresponding to the singular vectors $\underline{z}_i$ and $\underline{v}_i$. These signals represent only the portion of the interactive transmittances in the l^{th} loop which are scaled by σ_i, the direct transmittance in the l^{th} loop and the parallel transmittances for the other loops which pass through the i^{th} loop. For the 2×2 system shown in Figure 6.3, equation 6.1.23 gives,

$$\underline{y} = (\sigma_1 \underline{W}_1 + \sigma_2 \underline{W}_2)\underline{u} \qquad [6.1.56]$$

where,

$$\underline{W}_1 = \begin{bmatrix} z_{11}v_{11} & z_{11}v_{21} \\ z_{21}v_{11} & z_{21}v_{21} \end{bmatrix} \qquad \underline{W}_2 = \begin{bmatrix} z_{12}v_{12} & z_{12}v_{22} \\ z_{22}v_{12} & z_{22}v_{22} \end{bmatrix} \qquad [6.1.57]$$

Figures 6.6a and 6.6b show the signal transmittances which are represented by the rotational matrices $\underline{W}_1$ and $\underline{W}_2$. The elements of $\underline{W}_1$ indicate the relative magnitudes of the portion of the input signal $\underline{u}$ which is rotated to

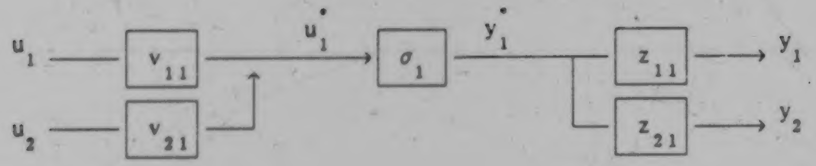
u_1
u_2
v_{11}
v_{21}
u_1^*
σ_1
y_1^*
z_{11}
z_{21}
y_1
y_2

Figure 6.6a: Signals Represented by $\underline{W}_1$

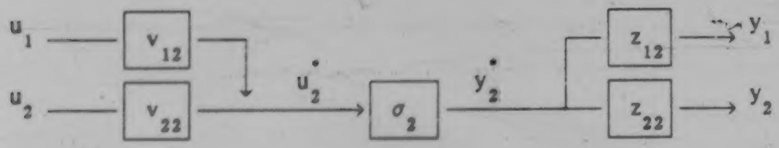
u_1
u_2
v_{12}
v_{22}
u_2^*
σ_2
y_2^*
z_{12}
z_{22}
y_1
y_2

Figure 6.6b: Signals Represented by $\underline{W}_2$

$\underline{v}_1$, $\dot{\underline{u}}_1$, is scaled by σ_1 and then reappears in the output direction $\underline{z}_1$, $\dot{\underline{y}}_1$. Similarly, $\underline{W}_2$ indicates the relative magnitudes of the portion of the input signal which is rotated by $\underline{v}_2$, $\dot{\underline{u}}_2$, is scaled by σ_2, and reappears in the output direction $\underline{z}_2$, $\dot{\underline{y}}_2$. Therefore, $\underline{W}_1$ and $\underline{W}_2$ do not contain all the information necessary to fully characterize the transmittances in loop 1 and loop 2, respectively.

For a 2×2 system, Lau's method will always give the same measure of interaction (or alignment) for both loops. If ψ and ϕ are the angles of rotation of the output space and the input space basis vectors, respectively, from the standard Euclidean vector space, $\underline{Z}$ and $\underline{V}$ can be written as,

$$\underline{Z} = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \qquad \underline{V} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \tag{6.1.58}$$

Only one angle is required to specify a rotation (or a vector position) in $\mathbb{R}^2$. The rotational matrices, $\underline{W}_i$, are given by,

$$\underline{W} = \begin{bmatrix} \cos\psi\cos\phi & \cos\psi\sin\phi \\ \sin\psi\cos\phi & \sin\psi\sin\phi \end{bmatrix} \qquad \underline{W} = \begin{bmatrix} \sin\psi\sin\phi & -\sin\psi\cos\phi \\ -\cos\psi\sin\phi & \cos\psi\cos\phi \end{bmatrix} \tag{6.1.59}$$

If the loop pairings are $u_i \rightarrow y_i$ $(i=1,2)$ for the i^{th} loop, then

$$\cos\Theta_1 = \cos\Theta_2 = |\cos\psi\cos\phi| \tag{6.1.60}$$

indicating that the interactions in both loops are equivalent. If the loop pairings are $u_i \rightarrow y_j$ $(i \neq j)$ for the i^{th} loop, then

$$\cos\Theta_1 = |\cos\psi\sin\phi| \qquad \cos\Theta_2 = |-\cos\psi\sin\phi| \tag{6.1.61}$$

Since $|\cos\psi\sin\phi| = |-\cos\psi\sin\phi|$, the interaction measure for both loops is the same.

For a 3×3 system, Lau's method will always give three interaction measures or alignment angles. In $\mathbb{R}^3$, three directional angles need to be specified for a rotation (or vector location). A vector, $\underline{x}$ $\mathbb{R}^3$, from the origin is defined by distances x_1, x_2, and x_3 from the $\underline{e}_2$-$\underline{e}_3$, $\underline{e}_1$-$\underline{e}_3$, and $\underline{e}_1$-$\underline{e}_2$ planes, respectively. The directional angles are given by,

$$\cos\alpha = \frac{x_1}{d} \qquad \cos\beta = \frac{x_2}{d} \qquad \cos\gamma = \frac{x_3}{d} \qquad [6.1.62]$$

where $d = \sqrt{x_1^2 + x_2^2 + x_3^2}$ which is the length of the vector. Although Lau's method gives three alignment angles,, the angles do not indicate the interactions in the three control loops.

Lau's interaction measure will only give an indication of the overall directional properties of the system and alert the designer to potential difficulties associated with decoupling. Lau's measure does not indicate process interactions in the individual control loops and does not indicate closed loop interactions. To examine process interactions, the alignment of both the input and the output space with the standard Euclidean space should be accounted for in a manner similar to the characteristic loci method. The direct and parallel transmittances, given in equation 6.1.21 should be compared to the interactive transmittances given in equation 6.1.22.

In the analysis of the propagation of an input signal through the process given in chapter 2.2.4, it was shown that the input signal entering the process is rotated by $\underline{V}^T$, scaled by $\underline{\Sigma}$, and rotated by $\underline{Z}$ before exiting the process. From equation 6.1.19, the effect of an input signal on the i^{th} output is given by

$$y_i = [i^{th} \text{ row } \underline{Z}]\cdot\underline{\Sigma}\cdot\underline{V}^T\cdot\underline{u} = \underline{f}_i\cdot\underline{u} \qquad [6.1.63]$$

Assuming that the optimum loop pairing for the i^{th} control loop is $u_i \rightarrow y_i$, the signal transmittances in the open loop are given by

$$y_i = \begin{bmatrix} z_{i1} \dots z_{ii} \dots z_{im} \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 \underline{v}_1^T \\ \vdots \\ \sigma_i \underline{v}_i^T \\ \vdots \\ \sigma_m \underline{v}_m^T \end{bmatrix} \cdot \underline{u} = \begin{bmatrix} \text{interactive} \\ \vdots \\ \text{direct+parallel} \\ \vdots \\ \text{interactive} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_m \end{bmatrix} \qquad [6.1.64]$$

The alignment angle $\phi_i(j\omega)$ of $\underline{f}_i(j\omega)$ with respect to $\underline{e}_i$ where

$$\cos\phi_i(j\omega) = \frac{|(\underline{f}_i(j\omega), \underline{e}_i)|}{\underline{f}_i(j\omega)} \qquad [6.1.65]$$

represents the ratio of the direct and parallel transmittances to the total open loop transmittance exiting from the ith control loop (y_i) and is a direct measure of process interactions. If

$$\underline{z}_i = \underline{v}_i = \underline{e}_i \qquad [6.1.66]$$

the input and the output space will be perfectly aligned with the standard Euclidean basis. In this case, interactive signal transmittances and parallel transmittances will not exist. As $\underline{z}_i$ ($\underline{v}_i$) and $\underline{e}_i$ become misaligned, the interactive signal transmittances in the ith loop will increase and the direct transmittance will decrease. However, parallel signal transmittances will exist only if $\underline{v}_i$ and $\underline{e}_i$ are also misaligned.

In the 2×2 system shown in Figure 6.3, the output can be expressed in terms of the standard basis vectors and the singular vectors of $\underline{Q}$ as,

$$\begin{aligned} y_1 &= \left\{ \sigma_1 \underline{e}_1^T \underline{z}_1 \underline{v}_1^+ \underline{e}_1 + \sigma_2 \underline{e}_1^T \underline{z}_2 \underline{v}_2^+ \underline{e}_1 \right\} u_1 + \left\{ \sigma_1 \underline{e}_1^T \underline{z}_1 \underline{v}_1^+ \underline{e}_2 + \sigma_2 \underline{e}_1^T \underline{z}_2 \underline{v}_2^+ \underline{e}_2 \right\} u_2 \\ &= \left\{ \sigma_1 z_{11} v_{11} + \sigma_2 z_{12} v_{12} \right\} u_1 + \left\{ \sigma_1 z_{11} v_{21} + \sigma_2 z_{12} v_{22} \right\} u_2 \qquad [6.1.67a] \end{aligned}$$

and

$$\begin{aligned} y_2 &= \left\{ \sigma_1 \underline{e}_2^T \underline{z}_1 \underline{v}_1^+ \underline{e}_1 + \sigma_2 \underline{e}_2^T \underline{z}_2 \underline{v}_2^+ \underline{e}_1 \right\} u_1 + \left\{ \sigma_1 \underline{e}_2^T \underline{z}_1 \underline{v}_1^+ \underline{e}_2 + \sigma_2 \underline{e}_2^T \underline{z}_2 \underline{v}_2^+ \underline{e}_2 \right\} u_2 \\ &= \left\{ \sigma_1 z_{21} v_{11} + \sigma_2 z_{22} v_{12} \right\} u_1 + \left\{ \sigma_1 z_{21} v_{21} + \sigma_2 z_{22} v_{22} \right\} u_2 \qquad [6.1.67b] \end{aligned}$$

For the output y_1,

$$\underline{f}_1 = \begin{bmatrix} z_{11} & z_{12} \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 v_{11} & \sigma_1 v_{21} \\ \sigma_2 v_{12} & \sigma_2 v_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1 z_{11} v_{11} + \sigma_2 z_{12} v_{12} \\ \sigma_1 z_{11} v_{21} + \sigma_2 z_{12} v_{22} \end{bmatrix}^T = \begin{bmatrix} f_{11} \\ f_{21} \end{bmatrix}^T$$

[6.1.68a]

Similarly, for an input y_2

$$\underline{f}_2 = \begin{bmatrix} z_{21} & z_{22} \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 v_{11} & \sigma_1 v_{21} \\ \sigma_2 v_{12} & \sigma_2 v_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1 z_{21} v_{11} + \sigma_2 z_{22} v_{12} \\ \sigma_1 z_{21} v_{21} + \sigma_2 z_{22} v_{22} \end{bmatrix}^T = \begin{bmatrix} f_{21} \\ f_{22} \end{bmatrix}^T$$

[6.1.68b]

The angle of alignment of $\underline{f}_1$ with the standard Euclidean vector $\underline{e}_1$ is

$$\cos\Theta_1 = \frac{f_{11}}{\underline{f}_1} \qquad [6.1.69]$$

If ϕ is the angle that $\underline{z}_1$ makes with the $\underline{e}_1$ axis and ψ is the angle that the scaled input $\sigma_1\underline{v}_1$ makes with the $\underline{e}_1$ axis,

$$\underline{f}_1 = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \cdot \begin{bmatrix} \cos\psi \\ \sin\psi \end{bmatrix}$$

$$\underline{f}_1 = \begin{bmatrix} \cos\phi\cos\psi - \sin\phi\sin\psi \\ \sin\phi\cos\psi + \cos\phi\sin\psi \end{bmatrix} = \begin{bmatrix} \cos(\phi+\psi) \\ \sin(\phi+\psi) \end{bmatrix} = \underline{f}_1 \cdot \begin{bmatrix} \cos\Theta \\ \sin\Theta_1 \end{bmatrix}$$

[6.1.70]

The relative magnitudes of the direct and parallel transmittance terms and the interactive terms are given by the cosine and the sine of the sum of the angles that $\sigma_1\underline{v}_1$ and $\underline{z}_1$ make with the $\underline{e}_1$ axis, respectively, scaled by the length of $\underline{f}_1$. Therefore, the open loop interactions and approximate first order closed loop interactions are determined by the misalignments with respect to the standard Euclidean space of both the input space ellipsoid and the output space ellipsoid expressed in terms of the $\underline{V}$ and $\underline{Z}$ bases. If the input space is aligned with the standard basis space, then the interactions will be completely determined by the misalignment of the output space from the standard space (and vice versa). If both the input and the output spaces are aligned with the standard basis vectors, then the system will be naturally decoupled and there will be no significant interactions in the system.

For higher order systems, the measure of alignment is given by the directional cosines of $\underline{f}_i$ with the $\underline{e}_i$ vector. For this case, the angle analysis gets more complicated because there are 'n' directional cosines required to define a single vector in n-dimensional space. In a 3×3 system, three directional cosines (or 2 angles in spherical coordinates) are required to define a principal axis direction of the input and output space ellipsoids. Using spherical coordinates, it can be shown that the angle of misalignment of the output space to the standard Euclidean vector space is determined by the angles that both the input and the output space make with the standard Euclidean basis vector.

The interaction measure based on the SVD of the process TFM, given in equation 6.1.65, is equivalent to the direct Nyquist array measure (Jensen, 1981) and the internal model control (IMC) interaction measure (Economou, 1986) Although the resulting interaction measures and the gains are plotted as a function of frequency in a Bode plot format, a geometric interpretation of the interaction is also available. If the TFM matrix is symmetric, the SVA and the CL method will produce the same geometric measure of interaction. Unlike Lau's measure, this measure will indicate the interactions in each loop because the misalignment of both the input space and the output space is accounted for.

Example 6.4: Interaction measures were calculated for the two interacting tank system given in the Appendix A for various values of the valve resistances. The objective of the system is to maintain the liquid level in the tank by manipulating the incoming flow streams. At steady state, the process transfer function matrix is given by

$$G_p(0) = \begin{bmatrix} R_1 + R_2 & R_2 \\ R_2 & R_2 \end{bmatrix} \qquad [6.1.71]$$

Figures 6.7 to 6.10 show the open-loop step responses to changes in the inputs for the four cases:

CASE	R_1	R_2
a	0.1	1.0
b	1.0	1.0
c	2.0	1.0
d	10.0	1.0

As R_2 increases with respect to R_1, the interactions exhibited in both control loops decrease with the interaction in the first loop decreasing significantly more than the second loop. From SVD of the transfer function matrix, the optimum input-output pairings are,

$$y_1 \Rightarrow u_1 \qquad [6.1.72a]$$

$$y_2 \Rightarrow u_2 \qquad [6.1.72b]$$

Figures 6.11 to 6.14 shows Lau's interaction measure, the alignment angles and the total interaction, for the four cases. In all cases, the alignment angles are the same for both loops and are equal to the total interaction measure. As the resistance of the middle valve increases, the alignment angles decrease indicating that the interactions in both loops are

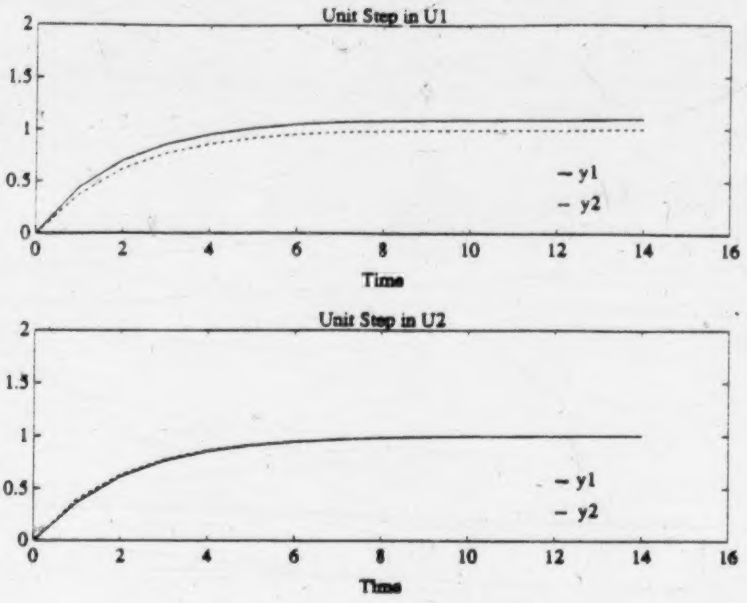

Figure 6.7: Open Loop Step Responses for Two Interacting Tank System (Case a)

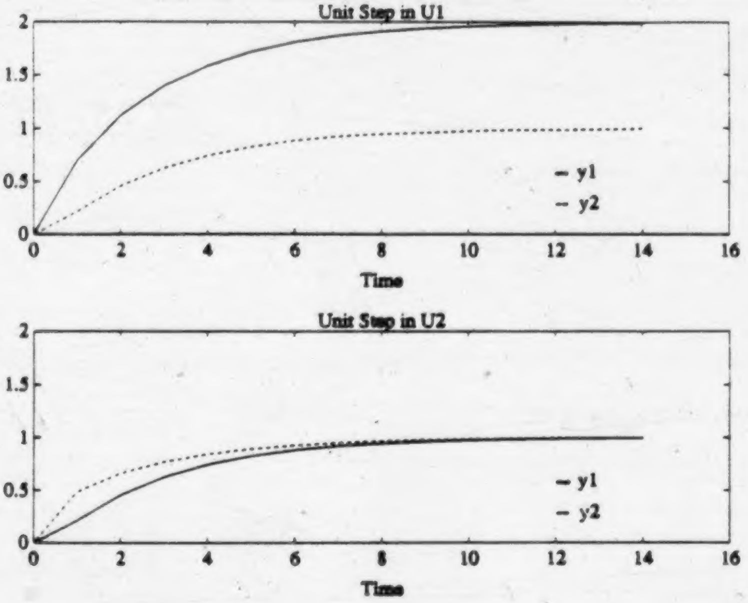

Figure 6.8: Open Loop Step Responses for Two Interacting Tank System (Case b)

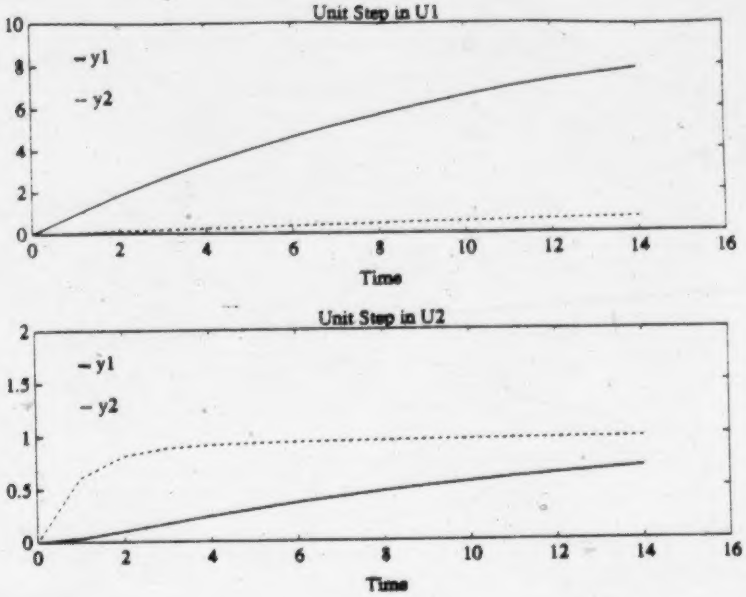

Figure 6.9: Open Loop Step Responses for Two Interacting Tank System (Case c)

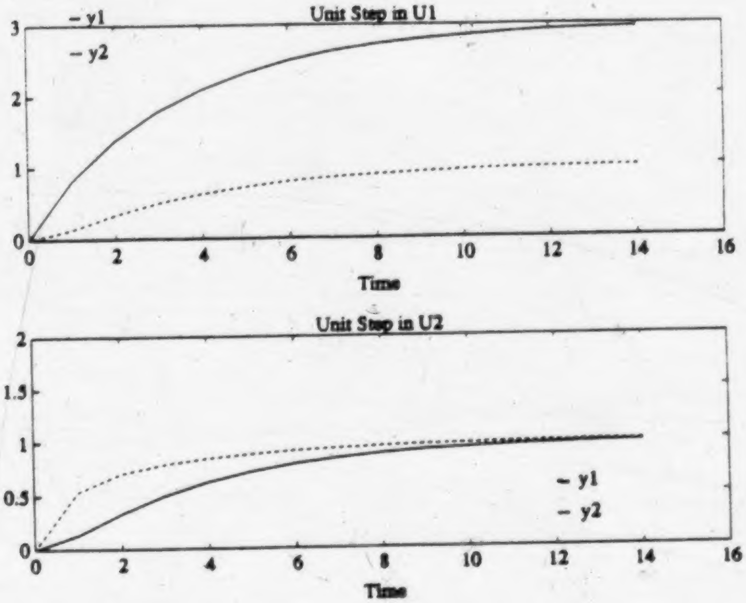

Figure 6.10: Open Loop Step Responses for Two Interacting Tank System (Case d)

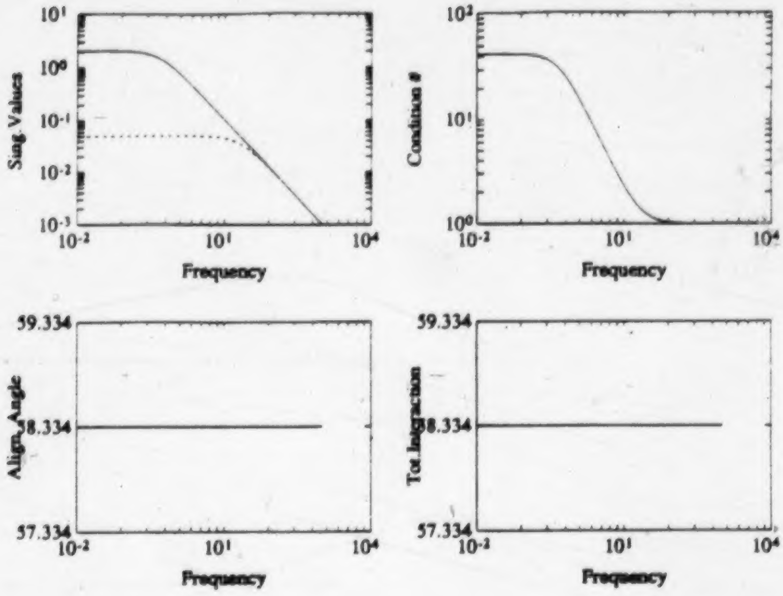

Figure 6.11: Lau's Measure of Interaction for Two Interacting Tank System (Case a)

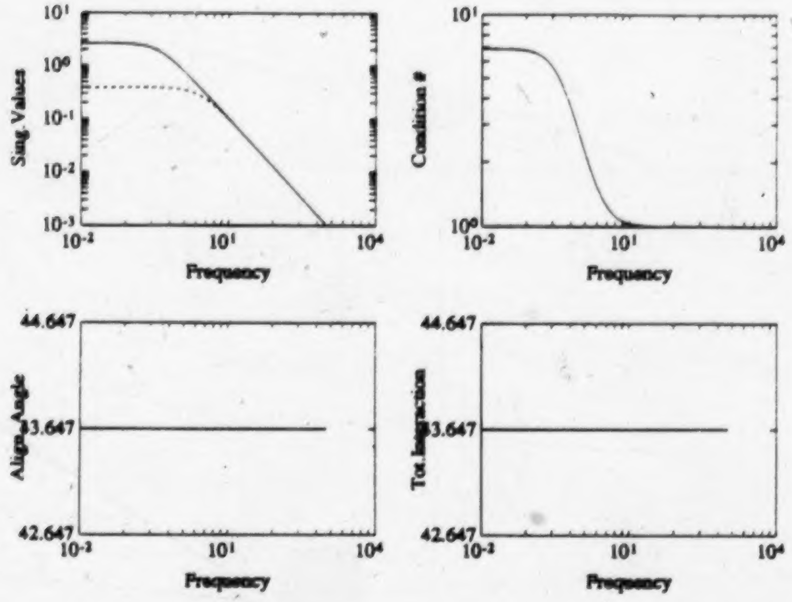

Figure 6.12: Lau's Measure of Interaction for Two Interacting Tank System (Case b)

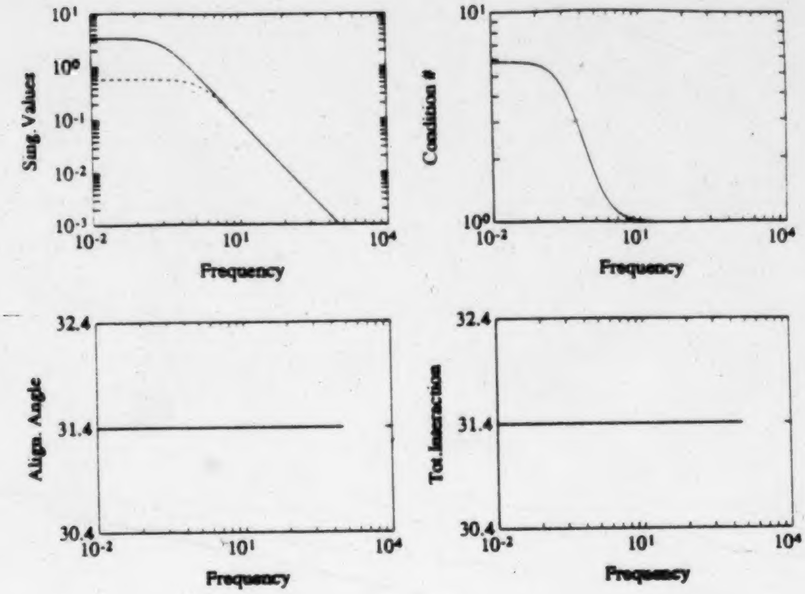

Figure 6.13: Lau's Measure of Interaction for Two Interacting Tank System (Case c)

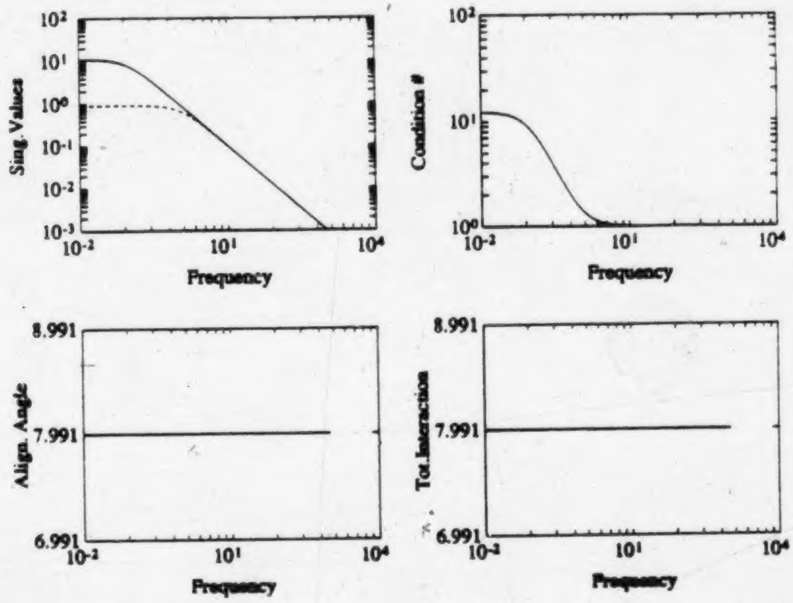

Figure 6.14: Lau's Measure of Interaction for Two Interacting Tank System (Case d)

decreasing and the system is becoming easier to decouple. However, the open loop step responses show that although the interaction in the first loop is decreasing significantly, the interaction in the second loop is significant. For the case where $R_1 = 1.0$ and $R_2 = 10.0$, the alignment angle is $<10^0$ indicating that there are no significant interactions in the system and the system is aligned very closely to the standard basis vectors. However, there is still significant interaction in the second loop which is not indicated by the interaction measure. Therefore, Lau's measure of interaction is indicating overall directional properties of the system but not the interactions between each input and output variable.

The DNA for the two interacting tanks, case a and case d, are shown in Figures 6.15 and 6.16, respectively. The DNA shows considerable interaction in both of the loops which is consistent with the open loop step responses shown in Figures 6.7 and 6.10. In addition, the DNA shows the interactions decreasing significantly in the first control loop as the resistances are changed. The dynamic RGA for case d is shown in Figure 6.17. The RGA indicates that there are no significant interactions in either loop. Therefore, the RGA provides the same information as the SVA interaction measure.

Example 6.5: The 2×2 system described by the process TFM (Jensen, 1986)

$$\underline{G}_p(s) = \begin{bmatrix} \dfrac{1}{(s+1)} & \dfrac{0.05}{(10s+1)} \\ \dfrac{1}{(2s+1)} & \dfrac{1}{(s+1)} \end{bmatrix} \qquad [6.1.73]$$

is approximately triangular. The optimum loop pairings obtained by matching the elements with the largest absolute values in each of the columns of $\underline{Z}$

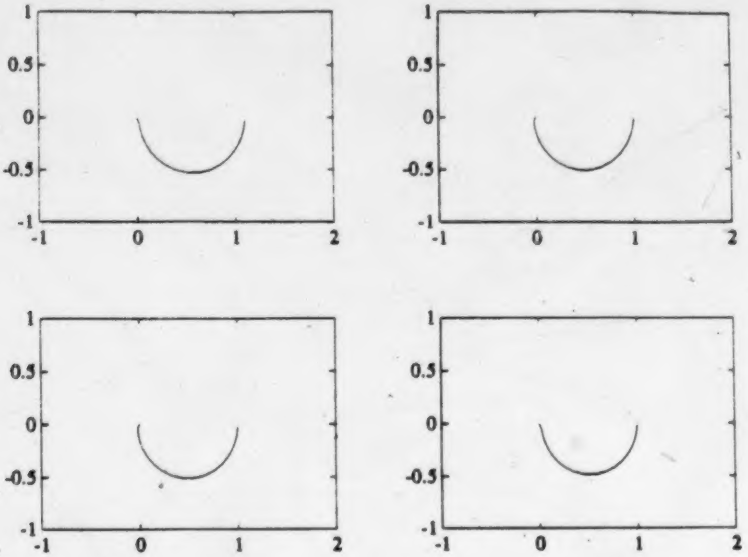

Figure 6.15: DNA for Two Interacting Tank System (Case a)

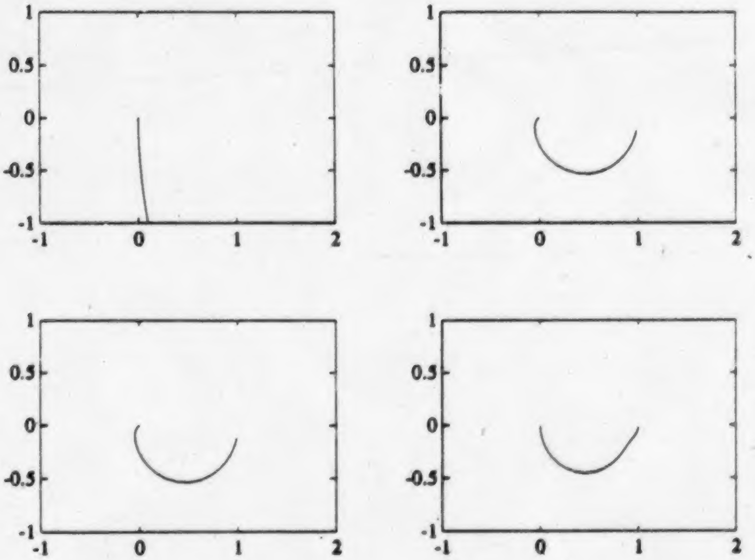

Figure 6.16: DNA for Two Interacting Tank System (Case d)

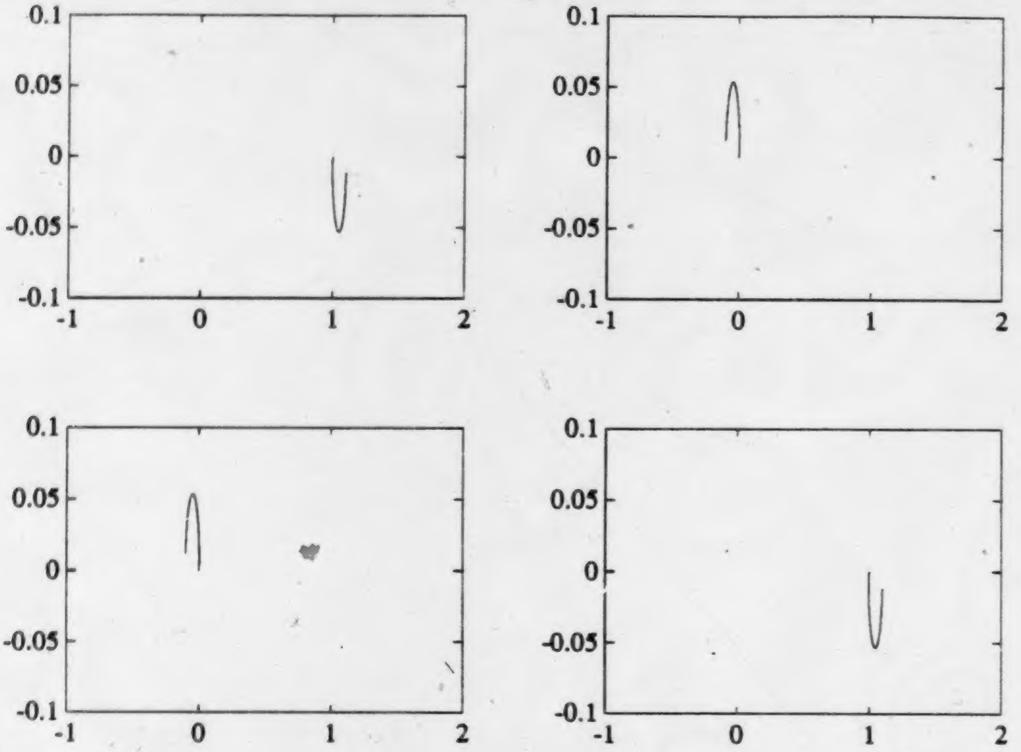

Figure 6.17: DRGA for Two Interacting Tank System (Case d)

and $\underline{Y}$, are

$$u_1 \Rightarrow y_1 \tag{6.1.74a}$$

$$u_2 \Rightarrow y_2 \tag{6.1.74b}$$

and are constant over the frequency range. The open loop step responses, given in Figure 6.18, show that the interactions in the first loop are small compared to those in the second loop. Figure 6.19 shows Lau's alignment angles and total interaction measure as a function of frequency for this process model. Although Lau's measure does indicate significant interaction in the system which may present problems in decoupler design, it fails to indicate the interactions between the inputs and the outputs in each of the control loops. Lau's interaction measures indicate that the interactions in both loops are equivalent and significant which is not the case. In a similar fashion, the (dynamic) RGA, shown in Figure 6.20, indicates that there is significant interaction in both control loops. The DNA is shown in Figure 6.21. Examination of the elements in the DNA show that although there is significant interaction in the second loop, the first loop is essentially non-interacting. Therefore, u_1 will affect only y_1 although u_2 will affect both y_1 and y_2. The open loop step responses justify this result.

Example 6.6: A distillation model with sidestreams (Lau, 1985a) was modeled using a 3×3 TFM. The compositions of the distillate (x_D) and the sidestreams (x_1 and x_2) are to be controlled by manipulating the reflux flow rate (R) and the draw off rate from each of the sidestreams (F_1 and F_2). The process TFM, $\underline{G}_p(s)$, is given by,

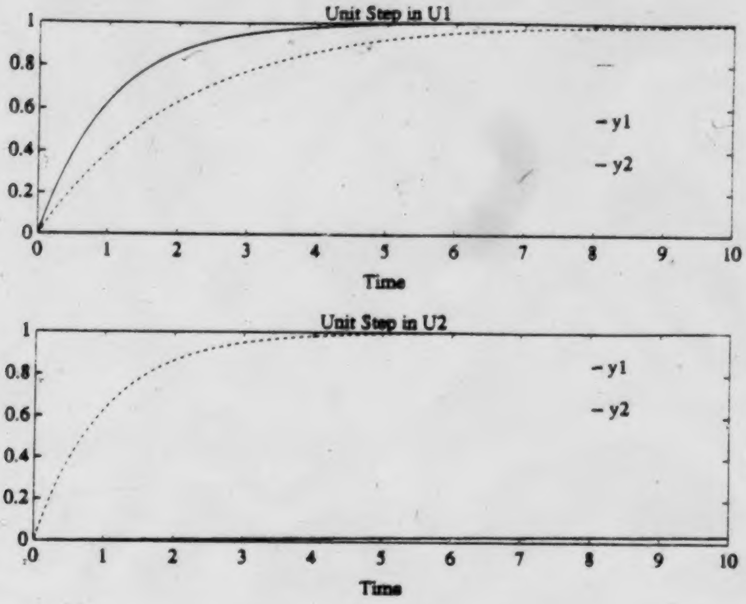

Figure 6.18: Open Loop Step Responses for Example 6.5 (Jensen, 1986)

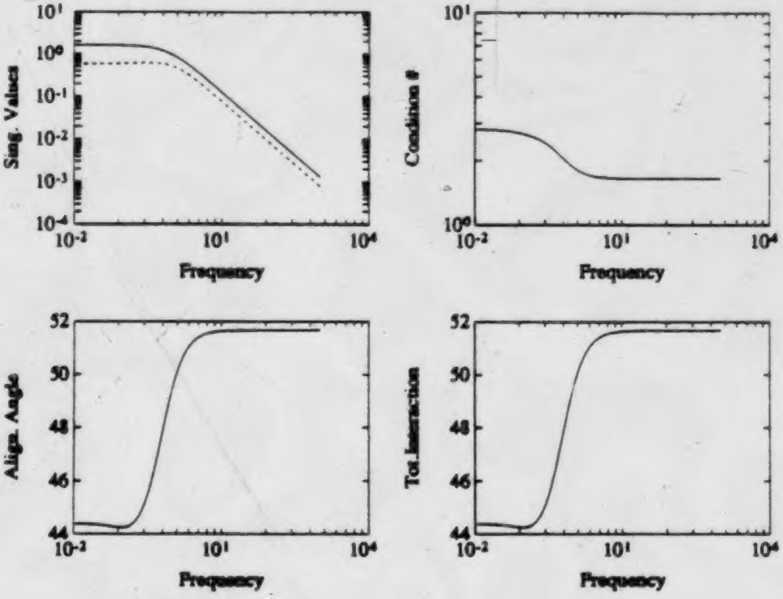

Figure 6.19: Lau's Measure of Interaction for Example 6.5 with Loop Pairings u_2-y_1 (σ_1) and u_1-y_2 (σ_2)

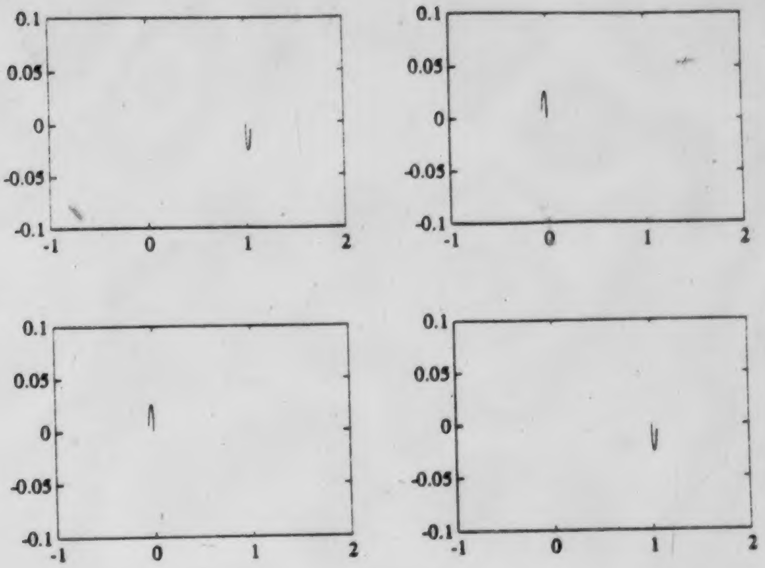

Figure 6.20: DRGA for Example 6.5 (Jensen, 1986)

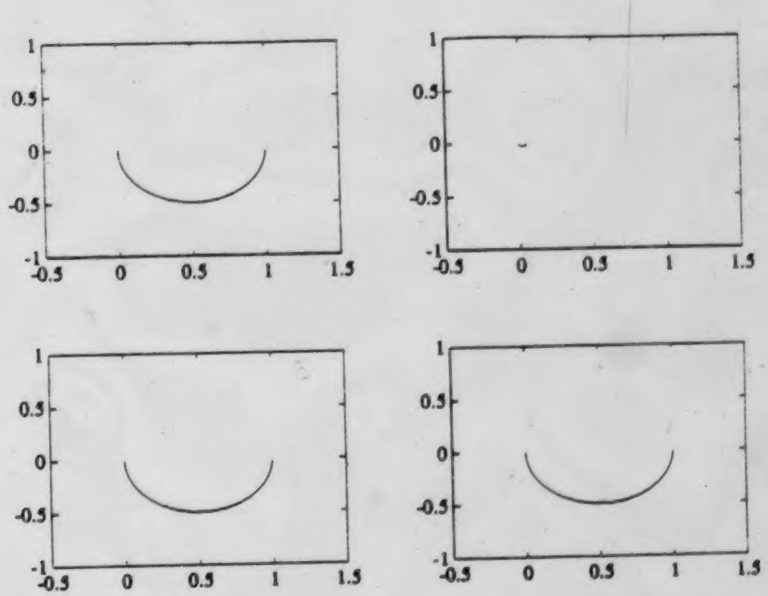

Figure 6.21: DNA for Example 6.5 (Jensen, 1986)

$$\begin{bmatrix} x_D \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{0.7}{(9s+1)} & 0 & 0 \\ \frac{2.0}{(8s+1)} & \frac{0.4}{(6s+1)} & 0 \\ \frac{2.3}{(10s+1)} & \frac{2.3}{(8s+1)} & \frac{2.1}{(7s+1)} \end{bmatrix} \cdot \begin{bmatrix} R \\ F_1 \\ F_2 \end{bmatrix} \qquad [6.1.75]$$

At steady state, the SVD components of $\underline{G}_p(0)$ are,

$$\underline{Z} = \begin{bmatrix} 0.118 & -0.337 & 0.934 \\ 0.389 & -0.850 & -0.356 \\ 0.914 & 0.405 & 0.031 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 4.189 & 0 & 0 \\ 0 & 1.443 & 0 \\ 0 & 0 & 0.097 \end{bmatrix} \qquad [6.1.76]$$

$$\underline{V} = \begin{bmatrix} 0.707 & -0.695 & 0.130 \\ 0.539 & 0.411 & -0.736 \\ 0.458 & 0.590 & 0.665 \end{bmatrix}$$

Matching the elements with the largest absolute values in the columns of $\underline{Z}$ and $\underline{V}$ result in loop pairings:

$$R \Rightarrow x_2 \qquad [6.1.77a]$$

$$F_2 \Rightarrow x_1 \qquad [6.1.77b]$$

$$F_1 \Rightarrow x_D \qquad [6.1.77c]$$

However, due to the triangular structure of the system, the loop pairings are constrained to be

$$R \Rightarrow x_D \qquad [6.1.78a]$$

$$F_1 \Rightarrow x_1 \qquad [6.1.78b]$$

$$F_2 \Rightarrow x_2 \qquad [6.1.78c]$$

to avoid controllability and observability problems. The open loop step responses, shown in Figure 6.22, indicate that there is significant interaction in the second and third loop but the first loop is effectively decoupled from the system. The singular values, Lau's alignment angles and

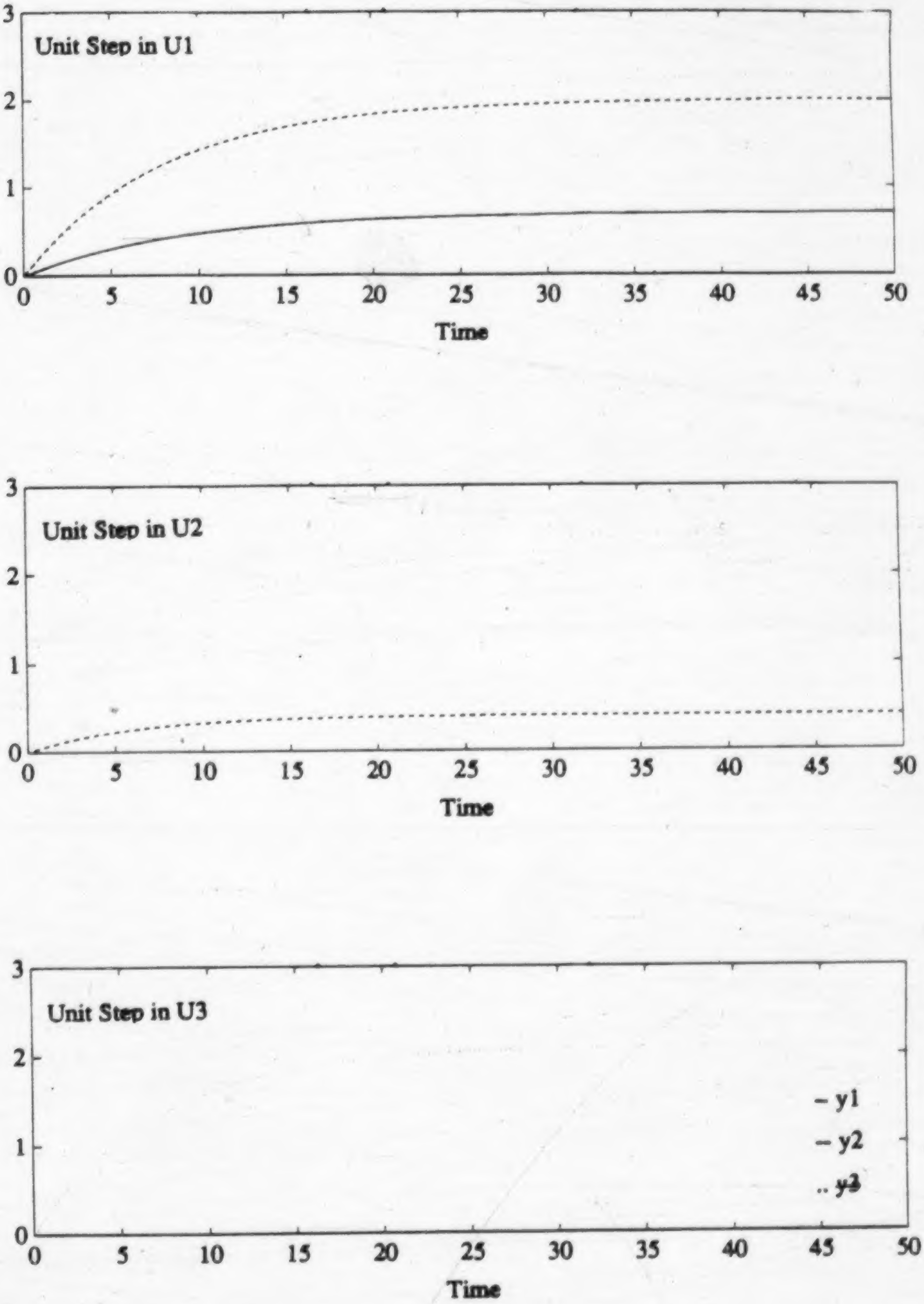

Figure 6.22: Open Loop Step Responses for Example 6.6 (Lau, 1985)

total interaction measure are shown as a function of frequency in Figure 6.23. The large alignment angles indicate that the system has very poor directional properties. However, Lau's interaction measure shows that there is significant interaction in all the loops with the third loop being slightly less. On the other hand, the DNA in Figure 6.24 indicates that the first loop has no interactions although there is interaction in the other two loops. Therefore, the DNA results are consistent with the step response information.

6.2 SVD in Non-Square System Analysis - Reduction in the Number of Controlled and/or Manipulated Variables

6.2.1 Optimum Sensor Location

In large systems where there are many possible manipulated variables and/or controlled variables, SVD of the process TFM can be used to reduce the system dimensions. The most common method for determining which sensors (or manipulated variables) to use in a control scheme is ad-hoc. Through simulation, the profiles of the measurement from each sensor is plotted before and after a given input sequence. The most sensitive sensor locations are determined from these profiles where the largest symmetrical deviation occurs. The optimum sensor locations (or controlled variables) are chosen to maximize the sensor (or controlled variable) sensitivity and minimize the process interactions. Maximizing the sensor (or controlled variable) sensitivity requires tightly tuned loops with high gains. Minimizing the process interactions corresponds to lower gains and more

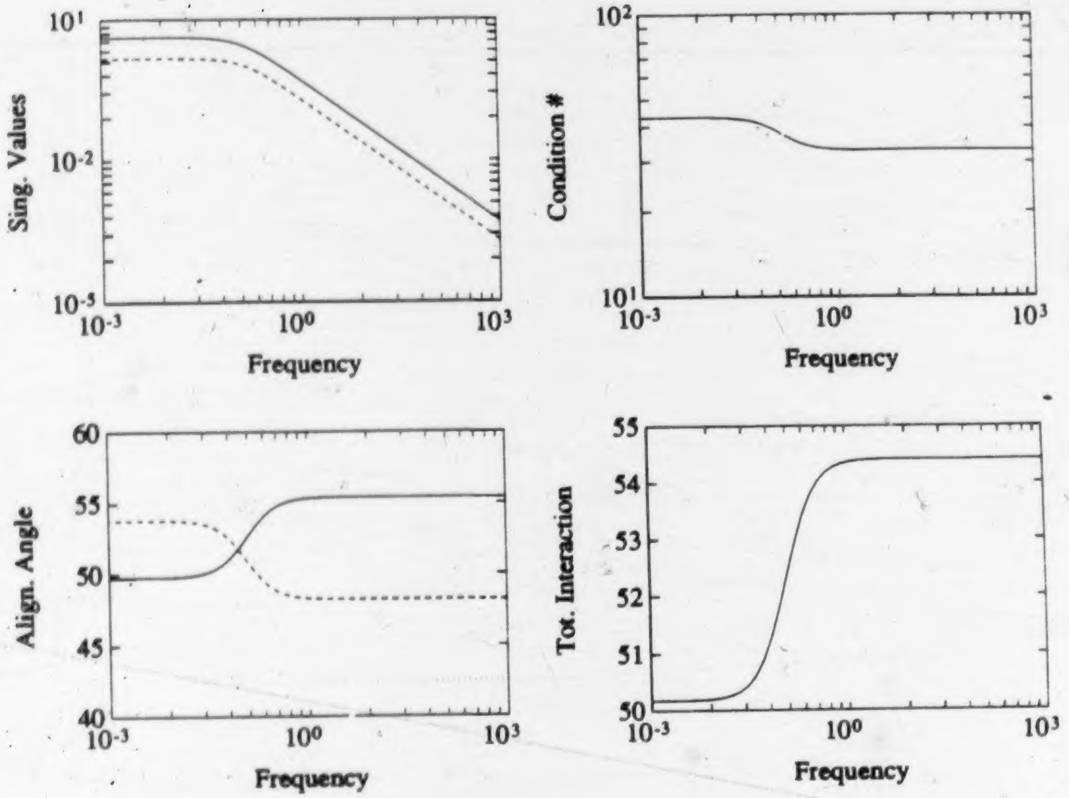

Figure 6.23: Lau's Measure of Interaction for Example 6.6 (Lau, 1985)

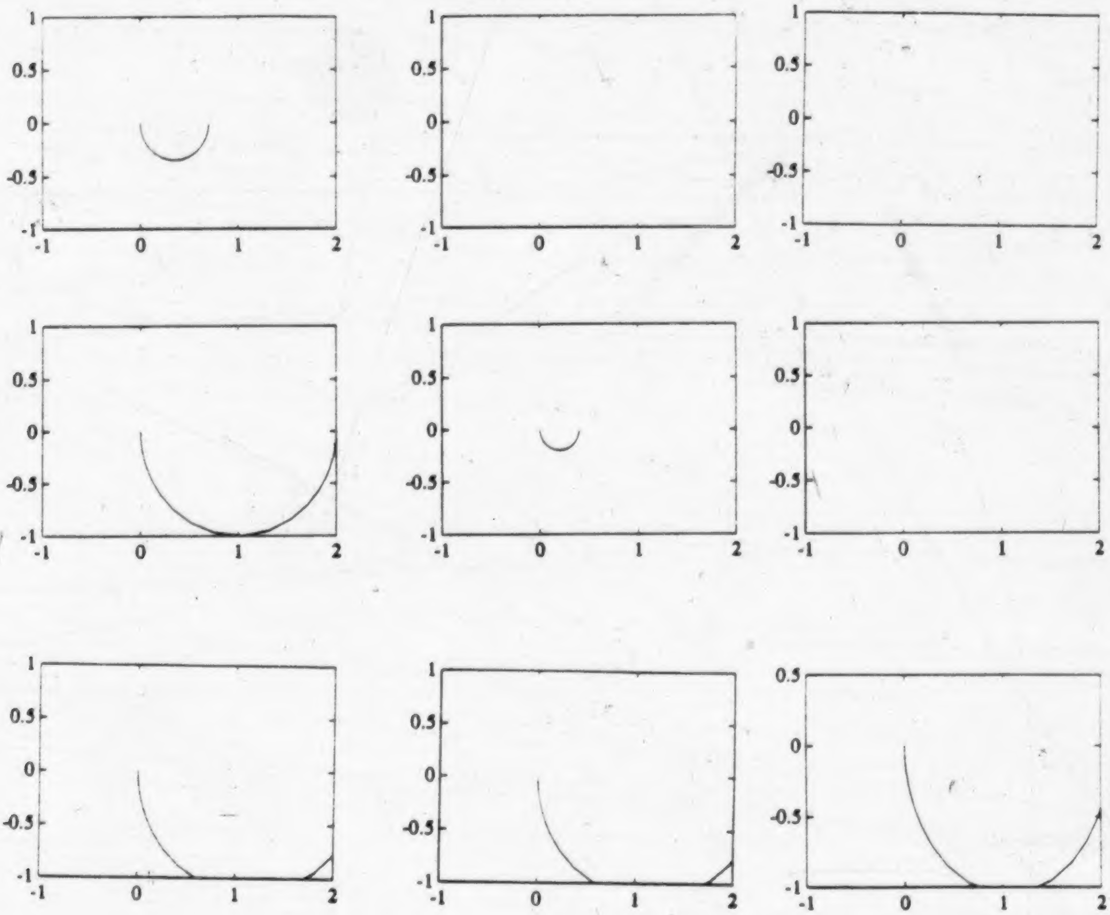

Figure 6.24: Direct Nyquist Array for Example 6.6

loosely tuned leading to a tradeoff in sensor location.

If the process transfer function matrix, $\underline{G}_p(s) \in \mathbb{R}^{m\text{x}n}$, is represented as,

$$\underline{G}_p(s) = \underline{Z}(s)\cdot\Sigma(s)\cdot\underline{V}^T(s) \qquad [6.2.1a]$$

where

$$\underline{\Sigma}(s) = \left(\begin{array}{c:c} \underline{S}(\bar{s}) & \underline{0} \\ \hline \underline{0} & \underline{0} \end{array}\right) \qquad [6.2.1b]$$

the sensor sensitivity (controlled variable sensitivity) can be maximized by choosing the sensors (controlled variables) corresponding to the most sensitive elements in the columns of $\underline{Z}$. If $\underline{G}_p(s)$ has rank r, which is less than m (full rank), (m-r) sensors and (n-r) manipulated variables can be deleted without changing the controllability or observability of the system (Lau, 1985a, Moore, 1986). If the condition number of $\underline{G}_p(s)$ is large, then the system is nearly singular and the rank of the system should be reduced to lower the condition number (Downs, 1981).

An example of a non-square system with more sensors than manipulative variables is a distillation column. In dual composition control of a distillation column, the output variables, usually the distillate and bottoms compositions, can be controlled from measurements of one or a combination of several tray temperatures (states) within the column by manipulating the reflux flow rate and the steam flow rate to the reboiler. Downs (1981), Moore (1986), and Bequette (1986) used a SVA approach to determine the optimum temperature sensor locations which balance the sensor sensitivity and interactions for a material balance control scheme. If the SVD components of the TFM relating the manipulated variables to the tray temperatures (states) is given in equation 6.2.1a, the projection of the temperature profile in the column onto $\underline{z}_1$ and $\underline{z}_2$ will result in two weighted average temperatures, $T_1 = \underline{z}_1\underline{T}$ and $T_2 = \underline{z}_2\underline{T}$. The weighted average temperatures could then be used in a control scheme instead of two tray temperatures. The average temperatures would include the effects of interactions and provide "optimal" sensitivity. However, a complete temperature profile is usually not available on most commercial distillation columns.

Usually, a material balance control scheme for a distillation column is implemented as a square system by choosing the two most sensitive trays as the measurable states or controlled variables. The relative magnitude of each element in $\underline{z}_1$ with respect to the other elements indicates the sensitivity of the corresponding sensor with respect to the manipulated variables (flow rates). The temperature of the trays corresponding to the largest elements in $\underline{z}_1$ and $\underline{z}_2$ are subject to the most change for a given input and are therefore the most sensitive. The relative impact of a tray on each loop (interaction) is indicated by the difference between the magnitudes of corresponding elements in $\underline{z}_1$ and $\underline{z}_2$. If $z_{i1} \simeq 1$, then $z_{i2} \simeq 0$ and the i^{th} sensor (i^{th} tray) will be affected by a change in u_1 only (non-interacting). Plots of $|z_{i1}|$ and $|\underline{z}_{i2}|$ versus the tray number or $|z_{i1}|-|z_{i2}|$ versus tray number will indicate both sensitivity and interaction of the sensors (Downs, 1981, Moore, 1986).

Example 6.7: The steady state TFM, $\underline{G}_p$, for a 50 tray ethanol-water distillation column (Moore, 1986) operating with a DQ control scheme (distillate rate and steam rate are the manipulated variables) is given in Table 6.1. The corresponding $\underline{Z}$, Σ, and $\underline{V}$ matrices are given in Table 6.2. A plot of $|\underline{z}_{i1}|$ and $|\underline{z}_{i2}|$ versus tray number is shown in Figure 6.25. Tray 34 and 44 correspond to the largest absolute values in $\underline{z}_1$ and $\underline{z}_2$, respectively. Therefore, these trays would be the most sensitive to an input. From the plot of $|\underline{z}_{i1}|$ - $|\underline{z}_{i2}|$ shown in Figure 6.26, the largest absolute values of the differences occur at trays 34 and 44. In this case, the most sensitive traysare also the least interactive. However, in most cases, the choice of the "optimum" sensor locations will involve a tradeoff between sensor sensitivity and interaction.

Bequette (1986) included the effects of disturbances when choosing the optimum sensor locations. If a state space model of the process is available, the sensitivity of the errors in the component compositions due to load changes under perfect measured variable control can be determined. The input, output and disturbance variables are expressed in deviation form such that $\underline{u} \leq 1.0$ and $\underline{d} \leq 1.0$. The state space model of a linear time invariant system is given by,

$$\dot{\underline{x}}(t) = \underline{A}\ \underline{x}(t) + \underline{B}\ \underline{u}(t) + \underline{E}\ \underline{d}(t) \qquad [6.2.2]$$

$$\underline{y}(t) = \underline{C}\ \underline{x}(t) \qquad [6.2.3]$$

$$\underline{x}_d(t) = \underline{S}\ \underline{x}(t) \qquad [6.2.4]$$

where $\underline{x}(t)$ is the vector of state variables, $\underline{d}(t)$ are the disturbances, $\underline{y}(t)$ is the vector of measured variables, $\underline{u}(t)$ is the vector of manipulated

Table 6.1 : Steady State TFM for the Ethanol-Water Column in Example 6.7

$$\begin{array}{r} 1\\ 2\\ 3\\ 4\\ 5\\ 6\\ 7\\ 8\\ 9\\ 10\\ 11\\ 12\\ 13\\ 14\\ 15\\ 16\\ 17\\ 18\\ 19\\ 20\\ 21\\ 22\\ 23\\ 24\\ 25\\ 26\\ 27\\ 28\\ 29\\ 30\\ 31\\ 32\\ 33\\ 34\\ 35\\ 36\\ 37\\ 38\\ 39\\ 40\\ 41\\ 42\\ 43\\ 44\\ 45\\ 46\\ 47\\ 48\\ 49\\ 50 \end{array} \quad \underline{G}_p = \left[\begin{array}{rr} 0.191e-03 & 0.000 \\ 0.191e-03 & 0.000 \\ 0.191e-03 & 0.000 \\ 0.191e-03 & 0.477e-04 \\ 0.191e-03 & 0.477e-04 \\ 0.191e-03 & 0.477e-04 \\ 0.334e-03 & 0.143e-03 \\ 0.381e-03 & 0.143e-03 \\ 0.381e-03 & 0.191e-03 \\ 0.429e-03 & 0.191e-03 \\ 0.572e-03 & 0.191e-03 \\ 0.572e-03 & 0.191e-03 \\ 0.763e-03 & 0.191e-03 \\ 0.858e-03 & 0.238e-03 \\ 0.954e-03 & 0.334e-03 \\ 0.114e-02 & 0.382e-03 \\ 0.134e-02 & 0.382e-03 \\ 0.162e-02 & 0.429e-03 \\ 0.191e-02 & 0.572e-03 \\ 0.234e-02 & 0.572e-03 \\ 0.277e-02 & 0.763e-03 \\ 0.334e-02 & 0.858e-03 \\ 0.401e-02 & 0.100e-02 \\ 0.496e-02 & 0.124e-02 \\ 0.610e-02 & 0.143e-02 \\ 0.758e-02 & 0.172e-02 \\ 0.954e-02 & 0.210e-02 \\ 0.121e-01 & 0.267e-02 \\ 0.157e-01 & 0.334e-02 \\ 0.212e-01 & 0.439e-02 \\ 0.299e-01 & 0.615e-02 \\ 0.452e-01 & 0.906e-02 \\ 0.685e-01 & 0.136e-01 \\ 0.823e-01 & 0.157e-01 \\ 0.222e-01 & 0.391e-02 \\ 0.238e-01 & 0.339e-02 \\ 0.252e-01 & 0.281e-02 \\ 0.261e-01 & 0.219e-02 \\ 0.264e-01 & 0.162e-02 \\ 0.261e-01 & 0.114e-02 \\ 0.254e-01 & 0.572e-03 \\ 0.240e-01 & 0.191e-03 \\ 0.222e-01 & -0.143e-03 \\ 0.200e-01 & -0.477e-03 \\ 0.176e-01 & -0.572e-03 \\ 0.152e-01 & -0.763e-03 \\ 0.128e-01 & -0.763e-03 \\ 0.105e-01 & -0.763e-03 \\ 0.839e-02 & -0.668e-03 \\ 0.324e-02 & -0.286e-03 \end{array}\right]$$

Table 6.2: SVD Matrices ($\underline{Z},\Sigma,\underline{V}$) for the Ethanol-Water Column in Example 6.7

$$\underline{Z} = \left[\begin{array}{r|rr}
1 & -0.125e-02 & -0.250e-02 \\
2 & -0.124e-02 & -0.250e-02 \\
3 & -0.125e-02 & -0.250e-02 \\
4 & -0.130e-02 & 0.159e-02 \\
5 & -0.130e-02 & 0.159e-02 \\
6 & -0.130e-02 & 0.159e-02 \\
7 & -0.233e-02 & 0.789e-02 \\
8 & -0.264e-02 & 0.727e-02 \\
9 & -0.269e-02 & 0.114e-01 \\
10 & -0.300e-02 & 0.107e-01 \\
11 & -0.394e-02 & 0.885e-02 \\
12 & -0.394e-02 & 0.885e-02 \\
13 & -0.519e-02 & 0.635e-02 \\
14 & -0.586e-02 & 0.919e-02 \\
15 & -0.658e-02 & 0.161e-01 \\
16 & -0.787e-02 & 0.177e-01 \\
17 & -0.912e-02 & 0.152e-01 \\
18 & -0.110e-01 & 0.155e-01 \\
19 & -0.131e-01 & 0.241e-01 \\
20 & -0.159e-01 & 0.184e-01 \\
21 & -0.189e-01 & 0.292e-01 \\
22 & -0.227e-01 & 0.298e-01 \\
23 & -0.272e-01 & 0.333e-01 \\
24 & -0.337e-01 & 0.413e-01 \\
25 & -0.414e-01 & 0.426e-01 \\
26 & -0.514e-01 & 0.478e-01 \\
27 & -0.645e-01 & 0.548e-01 \\
28 & -0.820e-01 & 0.701e-01 \\
29 & -0.106e+00 & 0.798e-01 \\
30 & -0.143e+00 & 0.985e-01 \\
31 & -0.202e+00 & 0.135e+00 \\
32 & -0.305e+00 & 0.185e+00 \\
33 & -0.462e+00 & 0.267e+00 \\
34 & -0.555e+00 & 0.269e+00 \\
35 & -0.149e+00 & 0.438e-01 \\
36 & -0.150e+00 & -0.225e-01 \\
37 & -0.168e+00 & -0.892e-01 \\
38 & -0.173e+00 & -0.154e+00 \\
39 & -0.175e+00 & -0.208e+00 \\
40 & -0.172e+00 & -0.245e+00 \\
41 & -0.167e+00 & -0.284e+00 \\
42 & -0.157e+00 & -0.299e+00 \\
43 & -0.145e+00 & -0.303e+00 \\
44 & -0.131e+00 & -0.304e+00 \\
45 & -0.115e+00 & -0.260e+00 \\
46 & -0.988e-01 & -0.265e+00 \\
47 & -0.832e-01 & -0.234e+00 \\
48 & -0.679e-01 & -0.203e+00 \\
49 & -0.543e-01 & -0.167e+00 \\
50 & -0.209e-01 & -0.671e-01
\end{array}\right]$$

$$\underline{V} = \begin{bmatrix} -0.988 & -0.151 \\ -0.151 & 0.988 \end{bmatrix}$$

$$\underline{\Sigma} = \begin{bmatrix} 0.151 & 0 \\ 0 & 0.012 \end{bmatrix}$$

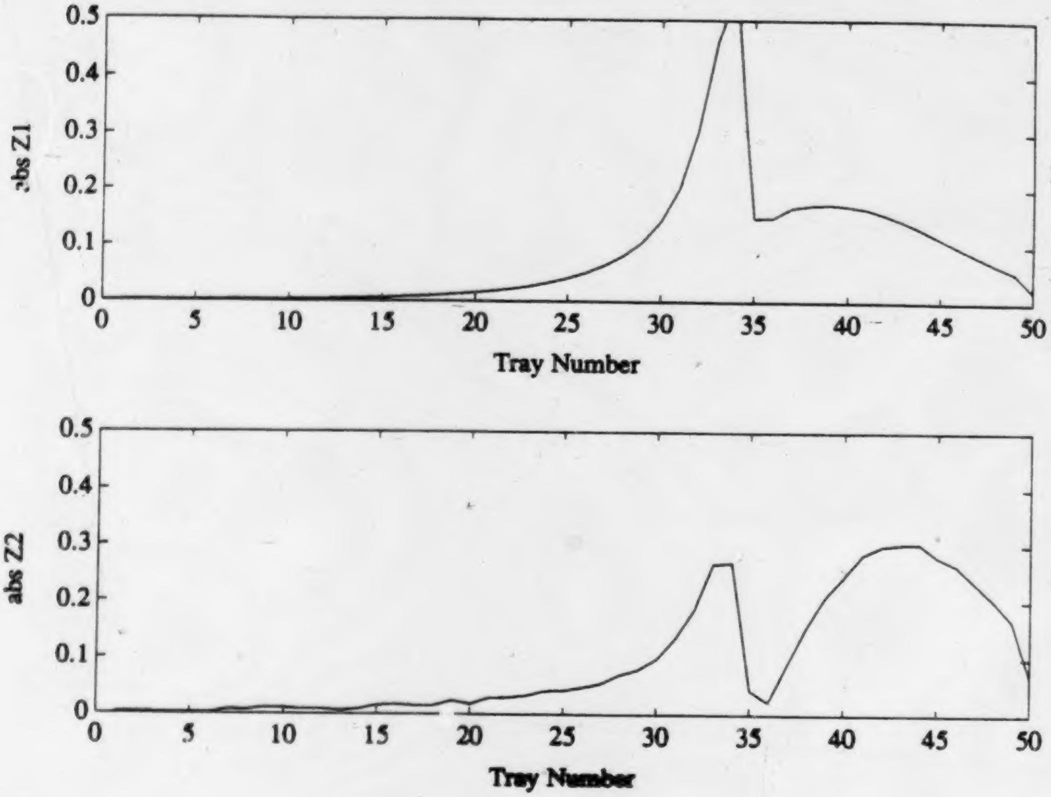

Figure 6.25: Plot of $|z_1|$ and $|z_2|$ Versus Tray Number for a 50 Tray Ethanol-Water Distillation Column in Example 6.7 (Moore,1986)

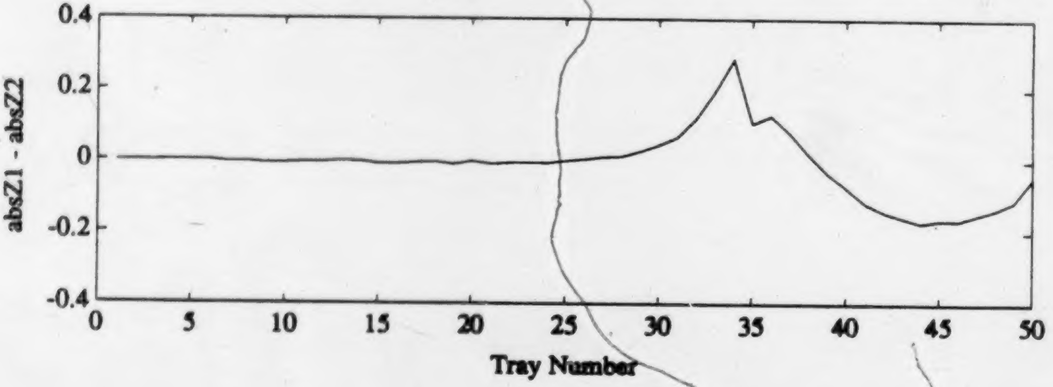

Figure 6.26: Plot of $|z_1|$ - $|z_2|$ Versus Tray Number for a 50 Tray Ethanol-Water Distillation Column in Example 6.7 (Moore, 1986)

variables and $\underline{x}_d(t)$ is the vector of controlled variables.

For the distillation column, the measured variables are the tray temperatures and

$$\underline{x}_d = \begin{bmatrix} y_D \\ x_B \end{bmatrix} \qquad \underline{u} = \begin{bmatrix} D \\ V \end{bmatrix} \quad \underline{d} = \begin{bmatrix} Fx_F q_F q_R \end{bmatrix}^T \qquad [6.2.5]$$

At steady state,

$$\dot{\underline{x}} = \underline{A}\,\underline{x} + \underline{B}\,\underline{u} + \underline{E}\,\underline{d} = 0 \qquad [6.2.6]$$

$$\therefore \underline{x} = -\underline{A}^{-1}\underline{B}\,\underline{u} - \underline{A}^{-1}\underline{E}\,\underline{d} \qquad [6.2.7]$$

assuming $\underline{A}$ is invertible. If the system is under perfect control with no offset,

$$\underline{y} = \underline{C}\,\underline{x} = -\underline{C}\,\underline{A}^{-1}\underline{B}\,\underline{u} - \underline{C}\,\underline{A}^{-1}\underline{E}\,\underline{d}$$

$$= \underline{C}\,\underline{G}_S\underline{u} + \underline{C}\,\underline{G}_{LS}\underline{d} = 0 \qquad [6.2.8a]$$

where

$$\underline{G}_S = -\underline{A}^{-1}\underline{B} = \text{state input gain matrix} \qquad [6.2.8b]$$

$$\underline{G}_{LS} = -\underline{A}^{-1}\underline{E} = \text{state load gain matrix} \qquad [6.2.8c]$$

The control action required to compensate for any disturbance, d(t), is

$$\underline{u} = -(\underline{C}\,\underline{G}_S)^{-1}\underline{C}\,\underline{G}_{LS}\underline{d} \qquad [6.2.9]$$

The error in the desired output compositions is,

$$\underline{x}_d = \underline{S}\,\underline{x} = \underline{S}\,(\underline{G}_S\,\underline{u} + \underline{G}_{LS}\underline{d})$$

$$= \underline{S}\,(\underline{G}_S\,(\underline{C}\,\underline{G}_S)^{-1}\underline{C}\,\underline{G}_{LS}\underline{d} + \underline{G}_{LS}\underline{d})$$

$$= \underline{S}\,(-\underline{G}_S\,(\underline{C}\,\underline{G}_S)^{-1}\underline{C}\,\underline{G}_{LS} + \underline{G}_{LS})\underline{d} \qquad [6.2.10a]$$

$$\therefore \underline{x}_d = \underline{G}_{LE}\,\underline{d} \qquad [6.2.10b]$$

where $\underline{G}_{LE} = \underline{S}\,(-\underline{G}_S\,(\underline{C}\,\underline{G}_S)^{-1}\underline{C}\,\underline{G}_{LS} + \underline{G}_{LS})$ [6.2.10c]

is the inferential error matrix. If $\underline{G}_{LE} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T$, $\underline{z}_1$ will indicate the strongest direction of composition deviation due to load changes whereas $\underline{v}_1$

will indicate the strongest disturbance direction. The deviation in the composition for any load disturbance $\underline{d}(t)$ should be as small as possible to ensure good regulatory control. Therefore,

$$\sigma_M(\underline{G}_{LE}) = \frac{\underline{x}_d}{\underline{d}} \qquad [6.2.11]$$

should be small for any $\underline{d}$. A load change where $\underline{d}$ = 1.0 in the worst direction, $\underline{v}_1$, would cause a steady state deviation in $\underline{x}_d$ of $\sigma_M(\underline{G}_{LE})$ under perfect feedback control. This represents a worst case scenario since $\underline{d} \leq 1.0$. Therefore, the choice of an optimum tray location should maximize the tray sensitivity and minimize the magnitude of the inferential error (maximum singular value of $\underline{G}_{LE}$) and the interactions.

6.2.2 Reduction of the Number of Manipulated Variables

For large systems with more possible manipulated variables than controlled variables, Keller (1987) used SVD in a modal method to reduce the number of inputs in the control scheme. The resulting system retained sufficient degrees of freedom to maintain the controllability. Other methods discussed in the literature are based on qualitative selection procedures and do not address the possible interactions within the system. Interactions between weak inputs and outputs (or state) can be poorly approximated in reduction methods because the strong inputs receive preferential treatment (Keller, 1987).

If a perfect process model is available, the control performance can be improved by increasing the number of manipulated variables (Morari, 1983).

However, in the presence of plant/model mismatch, the large number of manipulated variables leads to overparameterized controllers which may exhibit reduced robustness characteristics (Keller, 1987). In processes with a large number of inputs of various strengths, the controllers will compensate for weak inputs by increasing the control actions. In the presence of modeling error, the control actions can be applied in slightly different directions from that of the weak input and affect the stronger inputs. In addition, the more inputs incorporated into the control scheme, the higher the hardware costs and the more interactions which need to be described.

Let the linear time-invariant system be described by,

$$\dot{\underline{x}}(s) = \underline{A}\ \underline{x}(s) + \underline{B}\ \underline{u}(s) \qquad [6.2.12a]$$

$$\underline{y}(s) = \underline{C}\ \underline{x}(s) \qquad [6.2.12b]$$

which is assumed to be controllable and observable. The process variables are scaled such that they represent some physical quantity related to the process and thus can be compared numerically. Scaling of the process variables will be discussed in chapter 7. The transfer function associated with this state space model is,

$$\underline{G}(s) = \underline{C}(s\underline{I} - \underline{A})^{-1}\underline{B} \qquad [6.2.13]$$

A modal form of the system is obtained by a linear transformation of the original state space model such that,

$$\hat{\underline{C}} = \underline{C}\ \underline{T} \qquad \hat{\underline{B}} = \underline{T}^{-1}\underline{B} \qquad \hat{\underline{A}} = \underline{T}^{-1}\underline{A}\ \underline{T} \qquad [6.2.14a]$$

and

$$\hat{\underline{G}}(s) = \hat{\underline{C}}(s\underline{I} - \hat{\underline{A}})^{-1}\hat{\underline{B}} \qquad [6.1.14b]$$

The mode excitation, $\hat{\underline{B}}$, maps the input space $\mathbb{R}^n$ into the state space $\mathbb{R}^n$ and indicates the magnitude and direction of excitation from various inputs. The mode dynamics, $(s\underline{I}-\hat{\underline{A}})^{-1}$, maps $\mathbb{R}^n$ onto $\mathbb{R}^n$. The mode composition, $\hat{\underline{C}}$, maps $\mathbb{R}^n$ onto the output space. To determine the set of inputs which will be retained in the control scheme (strongest in terms of $\underline{G}(s)$), it is necessary to scale the modal space such that the dynamics and the compositions are the same for all modes. The mode dynamics are normalized to exhibit an unit steady state gain. The mode compositions ($\hat{C}$) are scaled so that they are equally well observed from $\underline{y}$. Keller (1987) suggested three methods: minimize the condition number of C, equalize the norms of the columns in C which will reduce the condition number significantly, or scale the modes to satisfy some output weighting criteria. The scaled mode excitation is given by a normalized $\tilde{\underline{B}}$ where each mode contributes the same maximal contribution to the dynamics of the system. If,

$$\tilde{\underline{B}} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T = \begin{bmatrix} \underline{Z}_1 & \underline{Z}_2 \end{bmatrix}\cdot\begin{bmatrix} \underline{\Sigma}_1 & \underline{0} \\ \underline{0} & \underline{\Sigma}_2 \end{bmatrix}\cdot\begin{bmatrix} \underline{V}_1^T \\ \underline{V}_2^T \end{bmatrix} \qquad [6.2.15]$$

the columns of $\underline{V}_1$ and $\underline{V}_2$ define an orthonormal coordinate system for viewing the inputs. If 'r' inputs are to be retained in the control scheme, $\underline{Z}$, $\underline{V}$, and $\underline{\Sigma}$ are partitioned such that $\underline{Z}_1$, $\underline{\Sigma}_1$ and $\underline{V}_1$ correspond to the retained (strongest) inputs. The inputs which are most strongly associated with $\underline{\Sigma}_1$ such that the largest proportion of the input signal is transmitted through $\underline{\Sigma}_1$ should be retained. The r inputs which are retained should be the most orthogonal to each other in their effect on the normalized modes. If

$$\underline{F} = \underline{v}_1^T\cdot\underline{v}_2^T \qquad [6.2.16]$$

f_{ii} is the fraction of the mode excitation originating from u_i and

f_{ii} is the fraction of the mode excitation originating from u_i and transmitted through $\underline{\Sigma}_1$. $(1-f_{ii})$ is the fraction of the mode excitation originating from u_i and transmitted through $\underline{\Sigma}_2$. Inputs corresponding to small diagonal elements, f_{ii}, are weak inputs and should be eliminated. The off-diagonal elements, f_{ij}, are a measure of the angle between the direction of excitation through u_i and u_j that goes through $\underline{\Sigma}_1$. If the off-diagonal elements in a row (or column) of $\underline{F}$ are zero, the corresponding $\underline{u}$ will excite the modes in a manner orthogonal to the other inputs. These inputs should be retained because no other inputs can replace this excitation. Enough inputs should be retained to ensure that each column of $\overline{\underline{V}_2}$ has at least one relatively large element corresponding to a retained input.

Johnston (1984, 1985) implemented internal control loops into a multiloop control schemes to decrease the sensitivity and minimize interactions in systems where there are more manipulated variables than measured variables. For a system described by a state space model where the process TFM is given by

$$\underline{G}_p(s) = \underline{C}(s\underline{I} - \underline{A})^{-1}\underline{B} \qquad [6.2.17]$$

The manipulated variables which minimize the condition number and maximize the magnitude of the smallest singular value over a wide frequency range are retained in the multiloop control scheme. The state input matrix, $\underline{B}$, is modified by the reduction in the number of manipulated variables. The condition number can be further improved by enhancing the dynamic relationships between the state variables by modifying the state matrix, $\underline{A}$. In systems with excess manipulated variables, this can be accomplished without major design changes. Excess manipulated variables can act as additional state variables by dynamically pairing with one or more existing (measurable) state and/or manipulated variables forming "internal control loops". The dynamic structure of the system has been modified by extending the state matrices. Careful selection of the internal control loop structure can enhance achievable control quality by reducing the condition number thereby reducing the sensitivity of the system to modeling errors. In addition, the choice of the optimum internal control loop structure should minimize both interactions and sensitivity over a wide frequency range. The optimum choice of measured variables in the internal control loop are the state variables which when held at their steady state values significantly improve the total interaction index (Johnston, 1985).

6.3 Measures of Controllability and Sensitivity of Multivariable Systems

SVA can be used to analyze the physical controllability of a system and sensitivity of a system to modeling errors independent of controller design. The condition number of a matrix, $\gamma(\underline{G})$ defined in equation 2.2.13, quantifies the sensitivity of the system with respect to uncertainties in the data or model (Lau, 1985a). If the condition number of the process TFM, $\gamma(\underline{G}_p)$, is large, $\underline{G}_p(s)$ is nearly singular (ill-conditioned) and small changes in the output variables, $\underline{y}$, or the model, $\underline{G}_p$, can result in large changes in the control actions, $\underline{u}$, which is generally undesirable in a real plant. Morari (1985) states that any multivariable controller provides an approximate inverse of the plant TFM. Therefore, the closer the approximation, the better the controller behaviour. If $\underline{G}_p(s)$ is ill-conditioned, it will be difficult to calculate the inverse of the process TFM and design a reliable decoupler, ideal or simplified, without numerical difficulties.

The condition number of the process TFM is the ratio of the maximum open loop decoupled gain to the minimum open-loop decoupled gain of a multiloop control scheme.

$$\gamma(\underline{G}_p) = \frac{\sigma_{max}(\underline{G}_p)}{\sigma_{min}(\underline{G}_p)} = \frac{\text{max.OL decoupled gain}}{\text{min.OL decoupled gain}} \qquad [6.3.1]$$

In an ill-conditioned system, the relative sensitivity of the system in one or more multivariable directions is weak. σ_1 is considerably larger than σ_n since the singular values are ordered in decreasing magnitude such that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. The ellipsoids representing the input and output spaces are very distorted. For a 2×2 system, if $\sigma_1 >> \sigma_2$,

$$\begin{aligned} y_1 &= (\sigma_1 z_{11} v^+_{11} + \sigma_2 z_{12} v^+_{12}) u_1 + (\sigma_1 z_{11} v^+_{21} + \sigma_2 z_{12} v^+_{22}) u_2 \\ &\simeq (\sigma_1 z_{11} v^+_{11}) u_1 + (\sigma_1 z_{11} v^+_{21}) u_2 \qquad [6.3.2] \end{aligned}$$

and

$$\begin{aligned} y_2 &= (\sigma_1 z_{21} v^+_{11} + \sigma_2 z_{22} v^+_{12}) u_1 + (\sigma_1 z_{21} v^+_{21} + \sigma_2 z_{22} v^+_{22}) u_2 \\ &\simeq (\sigma_1 z_{21} v^+_{11}) u_1 + (\sigma_1 z_{21} v^+_{21}) u_2 \qquad [6.3.3] \end{aligned}$$

The transmittances in the first loop are due to direct and interactive transmittances. The transmittances in the second loop are due entirely to interactive and parallel terms. Therefore, the system will be very difficult to decouple because there is no direct transmittance in the second loop.

The interpretation of controllability based on the condition number of the process TFM depends significantly on the process model (process TFM) and the scaling of the system variables. Scaling of the system will change the singular values of the process TFM and the condition number of the system. If the system is poorly scaled, cyclic responses may result which never reach a reasonable steady state (Moore, 1986). Scaling of the process variables will be discussed in chapter 7 of this thesis. M^cAvoy (1983), in an interaction analysis of steady state gain matrices, indicates that a condition number of 2 indicates a well behaved system with no decoupling problems. On the other hand, a condition number of 50 indicates a nearly singular system with decoupling problems. Maurath (1985) indicates that a condition number of 2000 for the dynamic matrix, formed from the step response coefficients (used in DMC) of a distillation column, indicates that the system is moderately ill-conditioned. Therefore, the condition number should be used only as a relative measure to choose between different control strategies with the same type of process model where all the process variables have been scaled in a similar manner. In systems with a large condition number, the process TFM is nearly singular indicating that one or more of the columns (or rows) are (nearly) linearly dependent. Therefore, all the control objectives can not be satisfied regardless of controller pairing or structure (Moore, 1981, Moore, 1986). The condition number of the system can be decreased by reducing the number of control objectives (controlled and/or manipulated variables).

In addition to the condition number, the magnitudes of the maximum and minimum singular values for a given system can be used to indicate controllability. Very small singular values correspond to weak decoupled process gains. These components are not very sensitive to control actions and require large controller gains. Therefore, large control actions are required despite a good condition number. In the presence of process constraints and noise, there could problems with the feedback scheme. On the other hand, very large singular values indicate that some components are very sensitive and require very small controller gains resulting in very small control actions which may be "lost" in the resolution of the manipulator.

The minimum singular value, $\sigma_{min}(\underline{G}_p)$, is a measure of the systems invertibility and indicates potential difficulties when implementing feedback control with inverse model based controllers (IMC, DMC, etc) and linear decouplers. If $\sigma_{min}(\underline{G}_p)$ is small, large control efforts will be required somewhere in the feedback scheme. Therefore, $\sigma_{min}(\underline{G}_p)$ should be large to obtain the "best" performance. Different control schemes can be compared based on their condition number and singular values to determine the best strategy to implement from a controllability viewpoint. Systems with small condition numbers and large minimum singular values will be the easiest to control.

Example 6.8: For the system of two interacting tanks given in example 6.4, the condition number of the system changes as the relative magnitude of the two resistances, R_1 and R_2 change. Therefore, the controllability of the system can be assessed based on the relative size of the valve resistances (or flows). Table 6.3 lists the condition number of $\underline{G}_p(0)$ and the singular values for various values of the valve resistances, R_1 and R_2.

Table 6.3: Condition Number and Singular Values for the Steady State Matrix of the Two Interacting Tank System

R_1	R_2	$\gamma(\underline{G})$	σ_i
0.1	1.0	42.076	2.051 0.049
0.5	1.0	10.404	2.281 0.219
1.0	1.0	6.854	2.618 0.382
2.0	1.0	5.828	3.414 0.585
5.0	1.0	7.670	6.193 0.807
10.0	1.0	12.319	11.099 0.901
1.0	10.0	42.076	20.513 0.488
1.0	5.0	22.155	10.525 0.475
1.0	2.0	10.404	4.562 0.438
1.0	0.1	12.319	1.110 0.090

If R_1 = 1.0 and R_2 = 1.0, the condition number of the system is 6.85. If R_1 = 0.1 and R_2 = 1.0, the condition number of the system is 42.08. Figures 6.27 and 6.28, respectively, show the propagation of an input signal through the ellipsoidal subspaces for these two cases. In Figure 6.28 where $R_1 \gg R_2$, the ellipsoid of the output space is very distorted. Therefore, a slight change in $\underline{u}$ results in a large change in $\underline{y}$ but only a slight change in direction. On the other hand, in Figure 6.27 where $R_1 = R_2$, a slight change in $\underline{u}$ will result in only a slight change in $\underline{y}$.

If $R_1 >> R_2$, the flow between the tanks, $q_1 = R_1/h_1$, will be small compared to the flow out of the second tank, q_2. The second tank acts effectively as a pipe and therefore, only the height in the first tank can be successfully controlled by q_0. The condition number of the system under these conditions is high indicating potential difficulties in satisfying both control objectives simultaneously. If $R_1 << R_2$, the flow between the tank is considerably more than the flow out of the tanks. The system will be difficult to control by manipulating both input flows independently because the tanks are effectively acting as one tank. Therefore, when R_1/R_2 is verylarge or very small, control is difficult (ie. system is "weakly controllable") as indicated by the relatively large amounts of manipulative action required to satisfy the control objectives. In these cases, it would be better to set one of the inlet flows to a constant rate then control the heights in the tanks by controlling the other flow. The "best" control situation is when the two resistances, R_1 and R_2 are of comparable magnitude. In this case, the condition number is low, $\sigma_{min}(\underline{G}_p)$ is maximum, and the interactions (shown in example 6.1) are minimal.

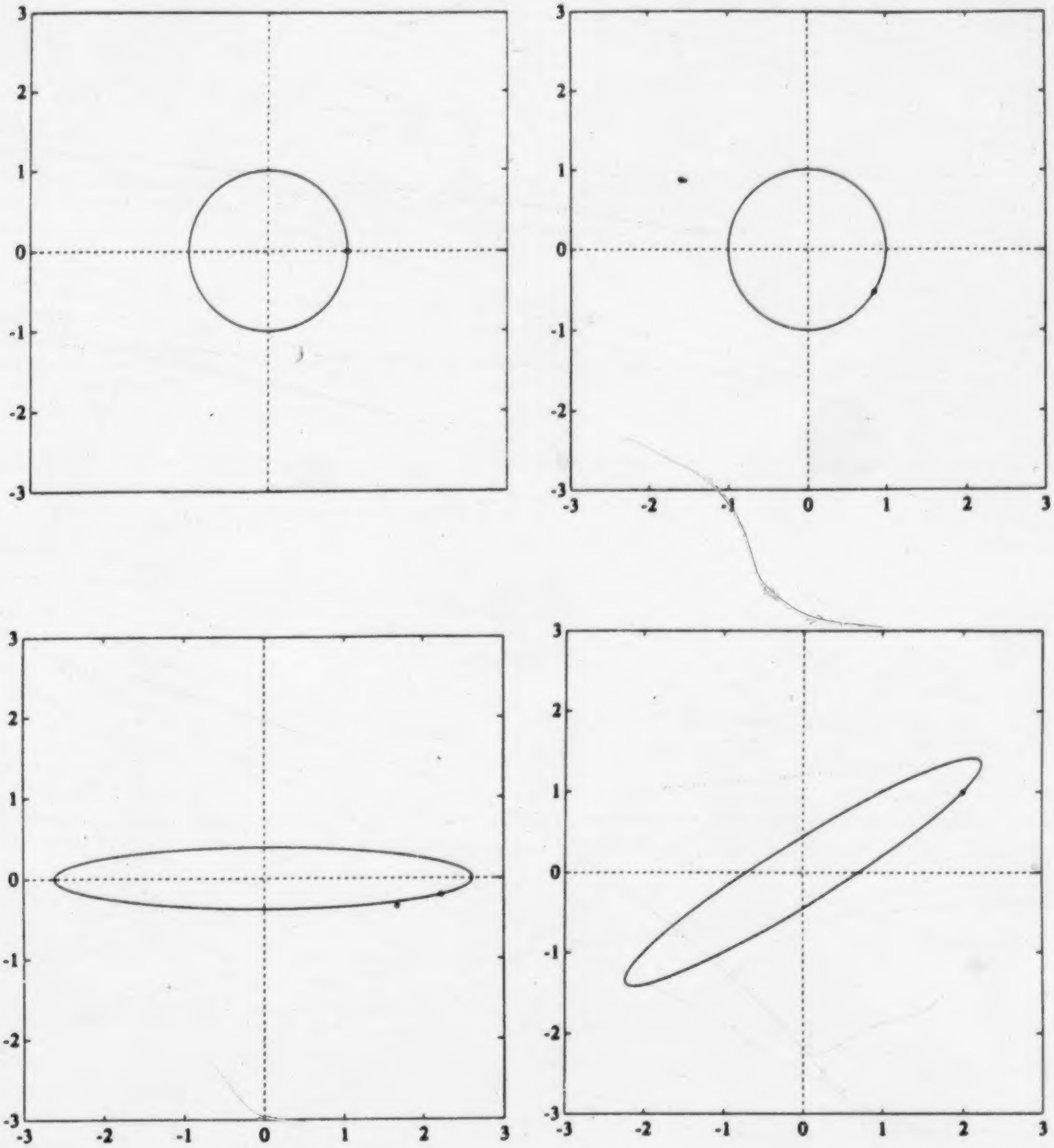

Figure 6.27: Input Signal Propagation Through the Two Interacting Tank System (Appendix A) with R_1=1.0 and R_2=1.0

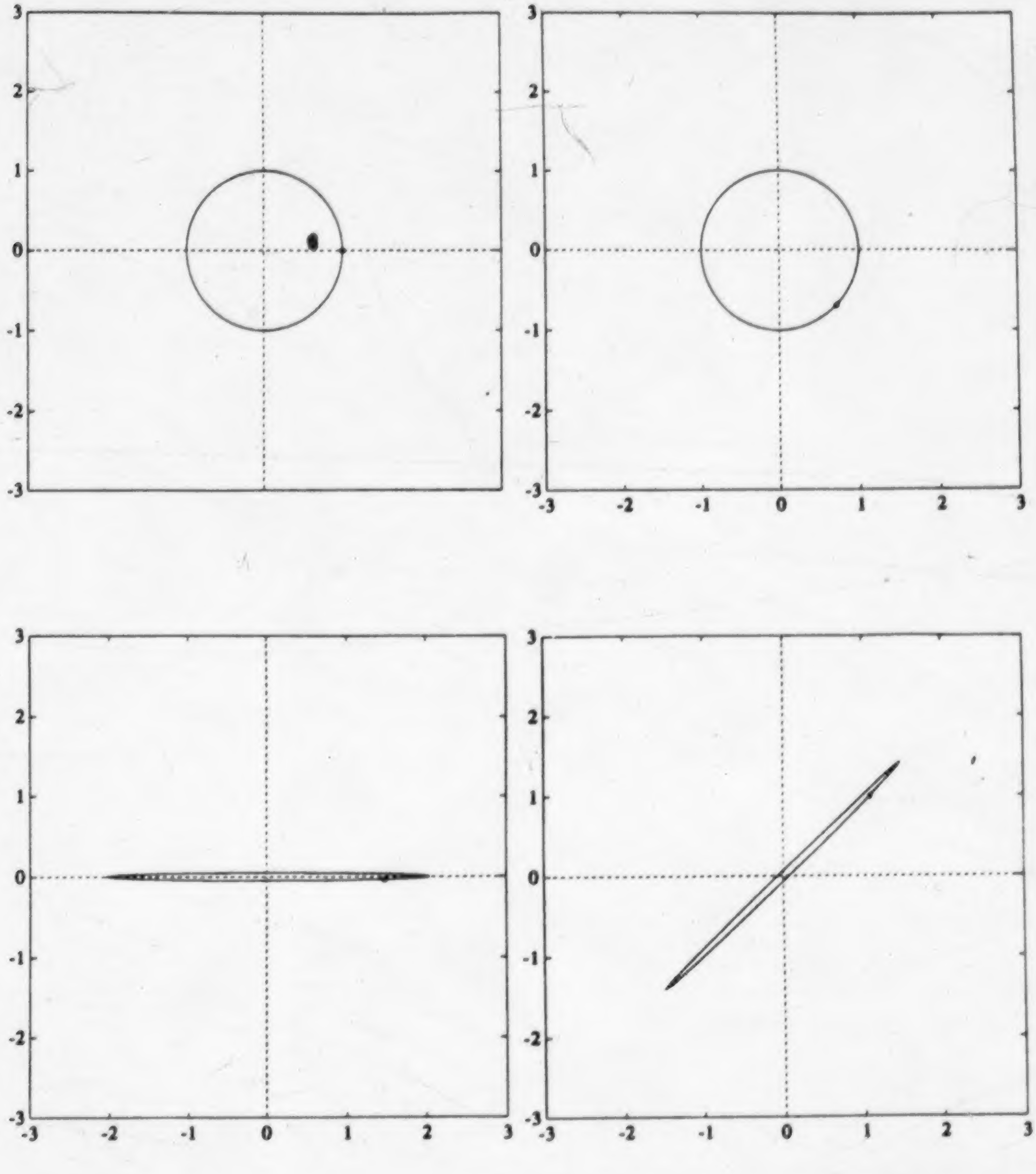

Figure 6.28: Input Signal Propagation Through the Two Interacting Tank System (Appendix A) with R_1=0.1 and R_2=1.0

6.4 Summary

Singular value decomposition of the system gain matrix provides a structural interpretation of process interactions. Signal transmittances which generate open loop and closed loop interactions between the signals entering the process and the resultant signals leaving the process can be"measured" directly using the singular vectors and their corresponding singular values. The projection of the singular vectors onto the input and output Euclidean vector space can always be used to indicate the degree ofinteraction in the system, since the singular vectors are always orthonormal. The direct Nyquist array interaction measure and the IMC interaction measure will produce the same results as the SVD measure. However, interaction measures such as Lau's which try to project the singular vectors onto a "combined input/output" vector space can lead to a false assessment of interactions. Although the measure may indicate there is considerable interaction in the system, it can not differentiate between the input signals and the output signals. Therefore, this type of measure could indicate that there is considerable interactions in both loops of a 2x2 system when there is considerable interaction only between one pair of input signals and output signals.

In the design of multiloop control systems, singular value decomposition of the system gain matrix provides a straightforward method of determining the optimum variable pairings for each control loop. The system can be analyzed dynamically to determine if the optimum variable pairings change over a given frequency range. This will assist the control engineer in evaluating the types of disturbances the control structure will be able to handle effectively with a given variable pairing.

The relative magnitudes of corresponding elements in the right and left singular vectors of the system gain matrix will indicate the relative sensitivity of that variable to changes in the process. Therefore, in systems which have more controllable variables than manipulative variables, SVD of the gain matrix can be used to design a square multiloop system which will maximize sensitivity and minimize interactions. As an alternative, linear combinations of the manipulated variables expressed in terms of the orthonormal basis sets $\underline{V}$ or controllable variables expressed in terms of the orthonormal basis set $\underline{Z}$ can be used in the multiloop strategy. This structure will inherently compensate for process interactions.

The controllability of the system can be analyzed from the singular value decomposition of the gain matrix independent of the controller structure. As the condition number of the system increases, the sensitivity of the system to noise and modelling errors increases which will decrease it's robustness. As the control system approaches singularity, it will become harder to control. The singular values of the gain matrix indicate the approximate process gain in the corresponding control loop. Therefore, the minimum singular value can be used to indicate the maximum controller gain required in the system. As the minimum singular value decreases, the control system will become harder to control because the process gain decreases.

All SVD analysis techniques are affected by the scaling of the process variables. Since scaling is the primary weakness of these methods, it is necessary that the system is scaled in an appropriate manner prior to application of the SVD techniques.

Chapter 7 Scaling of the Process Variables

7.0 Introduction

Scaling of the manipulated and/or controlled (state) variables is an important consideration in SVD analysis techniques. Scaling of a matrix involves pre and/or post multiplication of the matrix by diagonal matrices of scaling factors. Scaling can change the SVD components of a matrix (the singular values and vectors) and the condition number. Therefore, process scaling can affect the conclusions drawn from singular value analysis. If the scaling method utilized is inconsistent with physical limitations or constraints, the decomposition of the system matrices may not reflect the true characteristics of the system. The main objective of scaling should be to avoid domination of the results by some subset of the output, manipulated, or disturbance variables (Johnston, 1987).

The choice of a scaling procedure will depend on the primary use of the process model (Bonvin, 1985). However, any scaling method used to analyze the systems behaviour should eliminate the arbitrariness associated with the choice of physical units of the variables but maintain the relative strengths of the input-output (input-state) interactions. The scaling procedure should not alter the relative relationships between the inputs, outputs (states) and the disturbances. There are several scaling methods which have been recommended in the literature for use in the analysis of multivariable systems (Lau, 1985, Bonvin, 1985, Johnston, 1987, Keller, 1987). In the following sections, these methods are described and evaluated.

7.1 Scaling with Respect to a Reference Value

Scaling with respect to a reference value is a common scaling approach. The scaled variables are expressed as a percentage of the "full range" of the actual variable or as a deviation from their steady state (or nominal) values (Moore, 1981, M^{c}Avoy, 1983, Levien, 1985 Bequette, 1986, 1987 and Skogestad, 1987):

$$\tilde{y}_i = \frac{y_i^{max} - y_i}{y_i^{max} - y_i^{min}} \qquad \text{or} \qquad \tilde{y}_i = \frac{y_i - y_{ss}}{y_{ss}} \qquad [7.1.1]$$

where $\tilde{y}_i$ is the scaled variable. The ratio nature of the scaled variables eliminates the effects of the physical units making the scaled variables dimensionless. Scaling with respect to the maximum value accounts for the range of the variables. In this case, the controlled, manipulated and disturbance variables (y_i,u_i,d_i respectively) are scaled such that $0 \le |\tilde{u}_i| \le 1$, $0 \le |\tilde{y}_i| \le 1$ and $0 \le |\tilde{d}_i| \le 1$. This type of scaling is commonly referred to as "normalization".

Scaling with respect to a reference value could lead to resolution problems in a real system if the physical valves or sensors have been improperly sized. If the operating range of one of the variables is very large compared to the other variable, normalization or scaling in this manner will result in one of the variables having a very small scaled range whereas the other variable may be operating over the entire range 0 to 1. For example, if the operating range of a flow meter represents 10% of the full range of the meter and the operating range of another flow meter represents 80% of the full range, the flow meter which operates over the majority of the operating range will have more impact on the scaled system than the other regardless of the actual situation.

Example 7.1: Bonvin (1985, 1987) looked at a system of two continuous stirred tank heaters (CSTR) in series with first-order exothermic reactions, jacket or coil-cooling and a by-pass around the second reactor. The system was modeled from energy and mass conservation equations (see Appendix B). The process variables are scaled with respect to their steady state values. The resulting state space matrices ($\underline{A},\underline{B},\underline{C},\underline{D}$) and the corresponding steady state gain matrix , $\underline{G}_p(0)$, are given in Table 7.1. The condition number of the steady state gain matrix is 3.92×10^4. With the variables scaled with respect to their steady state values, the steady state system is ill-conditioned and control could be a problem in the presence of modeling errors and/or noise. Figure 7.1 shows the condition number and the singular values as functions of frequency. Although the condition number decreases at higher frequency, it is still relatively large.

7.2 Empirical Scaling Methods

There are several empirical scaling methods which have been used to scale variables for control system design. The manipulated and controlled variables can be scaled such that the maximum entry in each row and column of the process TFM is of the same magnitude. This scaling method is referred to as equilibration (Lau, 1985, Forsythe, 1977). Another empirical method is geometric scaling. Each row (column) in the matrix is scaled by the geometric mean of its largest and smallest elements. If $\underline{D}_r$ and $\underline{D}_c$ are

Table 7.1: State Space Realization (A,B,C,D) and Steady State Gain Matrix for the CSTR System in Example 7.1

$$\underline{A} = \begin{bmatrix} -3.18e\text{-}03 & 0 & -3.26e\text{-}05 & 0 \\ 1.00e\text{-}02 & -1.22e\text{-}02 & 0 & -2.69e\text{-}03 \\ 5.45e\text{-}02 & 0 & -3.10e\text{-}04 & 0 \\ 0 & 5.50e\text{-}02 & 1.00e\text{-}02 & -9.83e\text{-}03 \end{bmatrix}$$

$$\underline{B} = \begin{bmatrix} 6.86e\text{-}04 & 0 & 1.0e\text{-}03 & 0 & 0 & 0 \\ 5.70e\text{-}04 & 0 & 0 & 0 & 0 & -5.7e\text{-}03 \\ -1.41e\text{-}02 & 1.0e\text{-}03 & 0 & 1.25e\text{-}04 & 0 & 0 \\ -2.00e\text{-}03 & 0 & 0 & 0 & 5.0e\text{-}04 & -2.0e\text{-}02 \end{bmatrix}$$

$$\underline{C} = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 \\ 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 1.8 \end{bmatrix}$$

$$\underline{G}_P(0) = \begin{bmatrix} 0.243 & -0.012 & 0.112 & -0.00148 & 1.39e\text{-}15 & 1.23e\text{-}15 \\ 250.0 & -12.1 & 47.2 & -1.51 & -0.111 & 466.0 \\ -2.70 & 1.15 & 19.7 & 0.144 & -1.54e\text{-}16 & 1.54e\text{-}16 \\ -2.79 & 1.99 & 36.6 & 0.248 & 0.0905 & 1.03 \end{bmatrix}$$

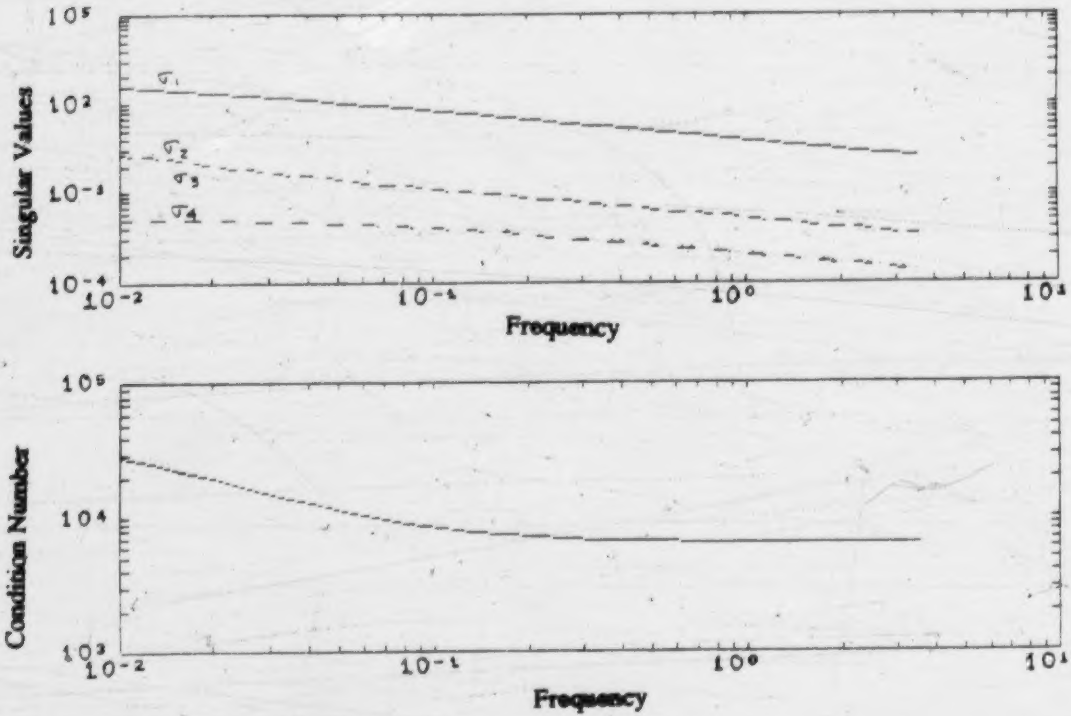

Figure 7.1: Condition Number and Singular Values of the Gain Matrix Versus Frequency for the CSTR System Scaled with Respect to Steady State Values (Bonvin, 1985, 1987)

the scaling matrices consisting of the row and column scaling factors, respectively, then

$$\underline{y} = \underline{D}_r^{-1}\underline{D}_r\underline{G}_p\underline{D}_c\underline{D}_c^{-1}\underline{u} \qquad [7.2.1]$$

$$\bar{\underline{y}} = \underline{D}_r\underline{y} = \underline{D}_r\underline{G}_p\underline{D}_c\underline{u} = \hat{\underline{G}}_p\bar{\underline{u}} \qquad [7.2.2]$$

Lau (1985) analyzed several different scaling techniques for a system matrix given by:

$$\underline{S} = \begin{bmatrix} \underline{A} & \underline{B} \\ \underline{C} & \underline{D} \end{bmatrix} \qquad \underline{S} = \begin{bmatrix} \underline{A}-s\underline{I} & \underline{B} \\ \underline{C} & \underline{D} \end{bmatrix} \qquad [7.2.3]$$

where the process model is described by a state space model. They found that empirical methods, such as equilibration and geometric scaling, could give different results depending on the order of scaling (row then column or vice versa).

Example 7.2: For the distillation column model given in example 6.2, the steady state gain matrix is

$$\underline{G}_p = \begin{bmatrix} 1.42 & -0.669 \\ 2.29 & -4.54 \end{bmatrix} \qquad [7.2.4]$$

If the matrix is scaled by the geometric mean of the rows followed by the geometric mean of the columns (geometric scaling), the resulting "scaled" matrix is

$$\hat{\underline{G}}_p = \begin{bmatrix} 1.432 & -0.698 \\ 0.698 & -1.432 \end{bmatrix} \qquad [7.2.5]$$

The scaling matrices used are

$$\underline{D}_r = \mathrm{diag}(0.975,3.224) \qquad [7.2.6]$$

$$\underline{D}_c = \mathrm{diag}(1.017.0.983) \qquad [7.2.7]$$

If the matrix is scaled by the geometric mean of the columns followed by the geometric mean of the rows, the resulting "scaled" matrix is the same as that given above. However, the scaling matrices are

$$\underline{D}_r = \mathrm{diag}\{0.550, 1.819\} \qquad [7.2.8]$$

$$\underline{D}_c = \mathrm{diag}\{1.803.1.743\} \qquad [7.2.9]$$

The scaled matrices were obtained using the geometric scaling routine outlined in Gill (1981). In this case, the order of scaling did not affect the resultant scaled matrix. The condition number of $\hat{\underline{G}}_p$ is 2.816 which is less than the condition number for $\underline{G}_p$ ($\gamma(\underline{G}_p) = 5.58$). The singular value decomposition components of the scaled matrix are

$$\hat{\underline{Z}} = \begin{bmatrix} 0.707 & -0.707 \\ 0.707 & 0.707 \end{bmatrix}$$

$$\hat{\underline{\Sigma}} = \begin{bmatrix} 2.131 & 0 \\ 0 & 0.734 \end{bmatrix}$$

$$\hat{\underline{V}} = \begin{bmatrix} 0.707 & -0.707 \\ -0.707 & -0.707 \end{bmatrix}$$

Example 7.3: For the model of the CSTR in example 7.1 (Bonvin, 1985), the steady state gain matrix when scaled by the geometric mean of the rows followed by that of the columns is given by

$$\underline{G}_{ss} = \begin{bmatrix} 1.4e{+}03 & -3.3e{+}02 & 3.9e{+}02 & -3.3e{+}02 & 5.7e{-}03 & 4.9e{-}04 \\ 2.5e{-}04 & -6.3e{-}05 & 3.0e{-}05 & -6.3e{-}05 & -8.3E{+}01 & 3.4e{+}04 \\ -7.4e{+}03 & 1.6e{+}04 & 3.4e{+}04 & 1.6e{+}04 & -3.1e{-}04 & 3.0e{-}05 \\ -6.0e{-}04 & 4.9e{-}04 & 1.1e{-}03 & 4.9e{-}04 & 3.2e{+}03 & 3.5e{+}03 \end{bmatrix}$$

The scaling matrices used are

$$\underline{D}_r = \mathrm{diag}\{5.959e{-}09, 32.64,\ 1.215e{-}08, 0.686\} \qquad [7.2.10]$$

$$\underline{D}_c = \mathrm{diag}\{3.007e{+}04, 5.931e{+}03, 4.845e{+}04, 741.3,$$
$$4.092e{-}05, 4.258e{-}04\} \qquad [7.2.11]$$

The condition number of the scaled matrix is 27.8 which is significantly lower than the gain matrix given in example 7.1. The singular value decomposition components are given in table 7.2. If the columns are scaled first followed by the rows, the resulting scaled steady state gain matrix will be

$$\underline{G}_{ss} = \begin{bmatrix} 1.5e{+}03 & -4.6e{+}02 & 5.4e{+}02 & -4.6e{+}02 & 6.2e{-}03 & 6.7e{-}04 \\ 2.0e{-}04 & -6.3e{-}05 & 3.0e{-}05 & -6.3e{-}05 & -6.6e{+}01 & 3.4e{+}04 \\ -5.9e{+}03 & 1.6e{+}04 & 3.4e{+}04 & 1.6e{+}04 & -2.5e{-}04 & 3.0e{-}05 \\ -1.7e{-}04 & 7.9e{-}04 & 1.8e{-}03 & 7.9e{-}04 & 4.2e{+}03 & 5.7e{+}03 \end{bmatrix}$$

The scaling matrices used are

$$\underline{D}_r = \mathrm{diag}(2.499e{-}05, 1.888e{+}05, 7.024e{-}05, 2.446e{+}03) \qquad [7.2.12]$$

$$\underline{D}_c = \mathrm{diag}(6.532, 1.026, 8.378, 0.128, 8.889e{-}09, 7.362e{-}08) \qquad [7.2.13]$$

The condition number of this scaled matrix is 23.9 which is slightly lower than the previous matrix. The SVD components of this matrix are shown in table 7.3. In this example, the order of scaling only slightly affected the resulting "scaled" matrix.

7.3 Normalization and Equilibration

Lau (1985b) recommends a scaling procedure involving normalization of the variables followed by formation of the matrix and row equilibration of the matrix. Normalization limits the variability in the system variables by removing the effects of physical dimensions. Row equilibration ensures that the maximum entry in each row has the same magnitude. Therefore, the effect of the largest normalized input on any given normalized output is of the same magnitude.

Table 7.2: SVD Components of the Scaled Steady State Gain Matrix for the CSTR System in Example 7.3 - Geometric Scaling (Row then Column)

$$\underline{Z} = \begin{bmatrix} 4.5e\text{-}03 & -1.6e\text{-}08 & 2.3e\text{-}06 & -1.0 \\ -3.2e\text{-}09 & -9.9e\text{-}01 & -1.1e\text{-}01 & -2.2e\text{-}07 \\ -1.0 & 6.5e\text{-}09 & -2.1e\text{-}08 & -4.5e\text{-}03 \\ -3.2e\text{-}08 & -1.1e\text{-}01 & 9.9e\text{-}01 & 2.2e\text{-}06 \end{bmatrix}$$

$$\underline{\Sigma} = \begin{bmatrix} 4.1e{+}03 & 0 & 0 & 0 \\ 0 & 3.4e{+}04 & 0 & 0 \\ 0 & 0 & 3.2e{+}03 & 0 \\ 0 & 0 & 0 & 1.5e{+}03 \end{bmatrix}$$

$$\underline{V} = \begin{bmatrix} 1.8e\text{-}01 & -9.1e\text{-}09 & 9.5e\text{-}07 & -9.0e\text{-}01 & -4.1e\text{-}01 & 0 \\ -3.9e\text{-}01 & 3.5e\text{-}09 & -1.9e\text{-}07 & 1.8e\text{-}01 & -5.6e\text{-}01 & -7.1e\text{-}01 \\ -8.2e\text{-}01 & 1.9e\text{-}09 & 3.9e\text{-}07 & -3.7e\text{-}01 & 4.5e\text{-}01 & -1.1e\text{-}06 \\ -3.9e\text{-}01 & 3.5e\text{-}09 & -1.9e\text{-}07 & 1.8e\text{-}01 & -5.6e\text{-}01 & -7.1e\text{-}01 \\ 5.7e\text{-}09 & -7.6e\text{-}03 & 1.0 & 1.1e\text{-}06 & 6.3e\text{-}11 & 4.2e\text{-}13 \\ -6.0e\text{-}09 & -1.0 & -7.6e\text{-}03 & 6.3e\text{-}10 & 5.7e\text{-}10 & 1.3e\text{-}14 \end{bmatrix}$$

Table 7.3 SVD Components of the Scaled Steady State Gain Matrix for the CSTR System in Example 7.3 - Geometric Scaling (Column then Row)

$$\underline{Z} = \begin{bmatrix} 3.3e\text{-}03 & -2.4e\text{-}08 & 1.8e\text{-}06 & -1.0 \\ -1.7e\text{-}08 & -9.9e\text{-}01 & -1.7e\text{-}01 & -2.9e\text{-}07 \\ -1.0 & 2.6e\text{-}08 & -4.5e\text{-}08 & -3.3e\text{-}03 \\ -5.5e\text{-}08 & -1.7e\text{-}01 & 9.9e\text{-}01 & 1.8e\text{-}06 \end{bmatrix}$$

$$\underline{\Sigma} = \begin{bmatrix} 4.1e{+}03 & 0 & 0 & 0 \\ 0 & 3.4e{+}04 & 0 & 0 \\ 0 & 0 & 4.1e{+}03 & 0 \\ 0 & 0 & 0 & 1.7e{+}03 \end{bmatrix}$$

$$\underline{V} = \begin{bmatrix} 1.4e\text{-}01 & -1.0e\text{-}08 & 6.8e\text{-}07 & -8.6e\text{-}01 & -4.9e\text{-}01 & 0 \\ -3.9e\text{-}01 & 1.0e\text{-}08 & -1.9e\text{-}07 & 2.4e\text{-}01 & -5.4e\text{-}01 & -7.1e\text{-}01 \\ -8.2e\text{-}01 & 1.5e\text{-}08 & 3.0e\text{-}07 & -3.8e\text{-}01 & 4.3e\text{-}01 & -1.1e\text{-}06 \\ -3.9e\text{-}01 & 1.0e\text{-}08 & -1.9e\text{-}07 & 2.4e\text{-}01 & -5.4e\text{-}01 & -7.1e\text{-}01 \\ 1.0e\text{-}09 & -1.9e\text{-}02 & 1.0 & 7.9e\text{-}07 & 7.5e\text{-}11 & 5.3e\text{-}13 \\ -2.2e\text{-}08 & -1.0 & -1.9e\text{-}02 & -6.9e\text{-}09 & 5.5e\text{-}10 & 1.3e\text{-}14 \end{bmatrix}$$

7.4 Minimization of Measures of Variability (Semi-Empirical Methods)

Lau (1985b) found that scaling methods which minimized a measure of variability in the entries of the process TFM did provide scaling factors with physical interpretations. A matrix is "well-scaled" if the variability of theentries in the matrix is minimal. Palazoglu (1985) used a multiobjective semi-infinite optimization method to determine the post and pre diagonal scaling matrices for a process TFM. The scaled process model was used in a robustness analysis of an IMC control structure. A scaling procedure was required such that the robustness indices and the upper bounds on their plant/model mismatch error did not depend on the units of the inputs and the outputs. Although this method may be optimal, the computations are very intense and may not be justified in a general multivariable analysis.

Morari (1985) suggests that the system be scaled such that the condition number of the steady state gain matrix, $\underline{G}_p(0)$, is minimized. In this case, the inputs and outputs are scaled such that the effect of the physical units are eliminated and the minimum and maximum singular values (decoupled loop gains) are of comparable magnitude. This type of scaling will ensure that the controller loop gains are comparable.

7.5 Scaling Methods Based on Physical Considerations

7.5.1 Dynamic Scaling Procedure (Bonvin, 1987)

Bonvin (1987) contends that scaling methods based on physical considerations are better than those based on empirical (or numerical) methods. They proposed an "ad-hoc" scaling method in which all the process equations are expressed on a common basis. The scaled process variables are all expressed in terms of the same quantity with the same physical units and therefore can be compared numerically. The relative magnitudes of the variables are not changed by the scaling procedure. If the magnitude of a variable is small before scaling, it will be relatively small after scaling. The scaling factors obtained are dependent on the steady state values of the variables. Therefore, the scaled process equations will be valid only in the vicinity of the operating point. If apriori knowledge of the energy and mass conservation equations for the process are unknown, this scaling procedure can not be used to scale the process transfer function matrices (or state matrices).

Each state variable, in a dynamic state space model of the process, can be described by a differential equation representing the conservation of some physical quantity (energy, mass, etc):

$$\dot{\underline{x}} = \underline{f}(\underline{x},\underline{u}) \qquad \underline{x}(0)=\underline{x}_0 \tag{7.5.1}$$

$$\underline{y} = \underline{g}(\underline{x}) \tag{7.5.2}$$

Let Π represent the physical quantity (energy or mass) which contains the highest level of information for the system:

$$\dot{\underline{\Pi}} = \underline{\hat{f}}(\underline{x},\underline{u}) \tag{7.5.3}$$

If all the states and the inputs are expressed in terms of $\underline{\Pi}$, they can be compared numerically since they have the same reference. The accumulation of each state, x_i, can be expressed in terms of an accumulation of Π by multiplying the state equation throughout by an equivalence factor, f_e^i, where

$$\Pi_i = f_e^i \cdot x_i \qquad [7.5.4]$$

A state scaling factor, p_i (i^{th} state) is defined such that each state, x_i, is expressed as an accumulation of Π.

$$\underline{\Pi} = \underline{\tilde{x}} = \underline{P}\underline{x} \qquad [7.5.5]$$

$$\underline{P} = \mathrm{diag}(p_i)\ ,\quad p_i > 0 \qquad [7.5.6]$$

Then,

$$\underline{\dot{\Pi}} = = \dot{\underline{\tilde{x}}} = \underline{\tilde{f}}(\underline{\Pi}, \underline{u}) = \underline{\tilde{f}}(\underline{\tilde{x}}, \underline{u}) \qquad [7.5.7]$$

$$\underline{y} = \underline{\tilde{g}}(\underline{\tilde{x}}) \qquad [7.5.8]$$

where $\dot{\Pi}_{si} = \dot{x}_{si} = \tilde{f}_{si}(\underline{\tilde{x}}, \underline{u})$ corresponds to the s^{th} state and the i^{th} input and $y_{js} = \tilde{g}_{js}(\tilde{x}_s)$ corresponds to the s^{th} state variable and the j^{th} output. The scaled states, $\underline{\tilde{x}}$, do not depend on the choice of units for the states, $\underline{x}$, although the scaling factors, p_i, do. The states are scaled such that a unit change in a scaled state variable corresponds to a unit change in Π.

The inputs are scaled such that each term involving an input represents a flow of Π into and out of the system. Input scaling factors, w_{si}, which relate changes in u_i to changes in the flow of Π into the subsystem associated with the s^{th} state variable are evaluated for the system at steady state

$$w_{si} = \left.\frac{\partial \tilde{f}_{si}}{\partial u_i}\right|_{\text{(at steady state)}} \qquad [7.5.8]$$

The input scaling factors, $w_{si,}$ are dependent only on the units of u_i because the states have already been normalized. Since each "state" equation may give a different scaling factor for each input, an "average" scaling factor for each input is defined. Bonvin chooses to use a scaling factor which is the maximum of all the individual scaling factors for a given input instead of an weighted average value. The input scaling factor for each input is

$$w_i = \max_s(|w_{si}|) \qquad [7.5.10]$$

and

$$\underline{\tilde{u}} = \underline{W}\cdot\underline{u} \qquad [7.5.11]$$

where $\underline{W} = \mathrm{diag}(w_i)$, $w_i > 0$. Each input is scaled such that a unit change in a scaled input brings one additional unit of Π into the strongest state variable.

Each output is expressed in terms of Π. A scaling factor, defined as

$$\frac{1}{v_{js}} = \left.\frac{\partial \tilde{g}_{js}}{\partial \underline{\tilde{x}}_s}\right|_{\text{(at steady state)}} \qquad [7.5.12]$$

relates changes in y_j to changes in $\tilde{x}_s$. An overall scaling factor, v_j, is determined for each output

$$v_j = \max_s \left(|v_{js}|\right) \qquad [7.5.13]$$

such that

$$\underline{\tilde{y}} = \underline{V}\underline{\tilde{x}} \qquad [7.5.14]$$

where $\underline{V} = \mathrm{diag}(v_j)$. Although the units of y_i affect the choice of v_{js}, $\tilde{y}_i$ is independent of the units of y_i. The outputs are scaled such that a unit change in a scaled output contains one unit of Π as information concerning the strongest state variable. There is no need with this scaling procedure to rescale the independent variable time. This feature is desirable because it leaves the eigenvalues of the system model unchanged (Bonvin, 1987).

Example 7.4: For the system of two continuous stirred tank heaters (CSTR) in series in example 7.1, energy can be chosen as the reference quantity, Π. Bonvin (1987) evaluated the model, given by conservation equations, using numerical data. A scaled model

$$\dot{\underline{x}}(0) = \hat{\underline{A}}\underline{x}(0) + \hat{\underline{B}}\underline{u}(0) \qquad [7.5.15]$$

$$\underline{y}(0) = \hat{\underline{C}}\underline{x}(0) \qquad [7.5.16]$$

where $(\hat{A},\hat{B},\hat{C})$ are

$$\hat{\underline{A}} = \begin{bmatrix} -3.18e-03 & 0 & -8.15e-04 & 0 \\ 1.00e-04 & -1.22e-02 & 0 & -6.73e-04 \\ 2.18e-03 & 0 & -3.10e-04 & 0 \\ 0 & 2.20e-03 & 1.00e-04 & -9.83e-03 \end{bmatrix}$$

$$\hat{\underline{B}} = \begin{bmatrix} 1.0 & 0 & 1.0 & 0 & 0 & 0 \\ 8.31e-03 & 0 & 0 & 0 & 0 & 1.0 \\ -8.22e-01 & 1.0 & 0 & 1.0 & 0 & 0 \\ -1.17e-03 & 0 & 0 & 0 & 1.0 & -1.4e-01 \end{bmatrix}$$

$$\hat{\underline{C}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\underline{G}\ (0) = \begin{bmatrix} 354.8 & -29.51 & 112.3 & -295.1 & 1.85e-15 & 0 \\ 3.639 & -3.028 & 0.471 & -3.028 & -5.545 & 81.77 \\ -157.1 & 1151.4 & 789.4 & 1151.4 & 2.47e-15 & 3.39e-15 \\ -0.903 & 11.04 & 8.136 & 11.036 & 100.5 & 4.053 \end{bmatrix}$$

was obtained using the scaling factors,

$$\underline{P} = \mathrm{diag}\left(10^8,\ 10^6,\ 4.0\times10^6,\ 4.0\times10^4 \right) \qquad [7.5.17]$$

$$\underline{W} = \mathrm{diag}\left(6.86\times10^4,\ 4000,\ 10^5,\ 500,\ 20,\ 5.7\times10^3 \right) \qquad [7.5.18]$$

$$\underline{Y} = \mathrm{diag}\left(10^8,\ 10^3,\ 4.0\times10^6,\ 2.22\times10^4 \right) \qquad [7.2.19]$$

The condition number of the original steady state gain matrix, $\underline{G}_p(0)$ is 3.92×10^4. The condition number of the scaled model, $\hat{\underline{G}}_p(0)$, is 22.3. Figure 7.2 shows the variation in the condition number and the singular values as

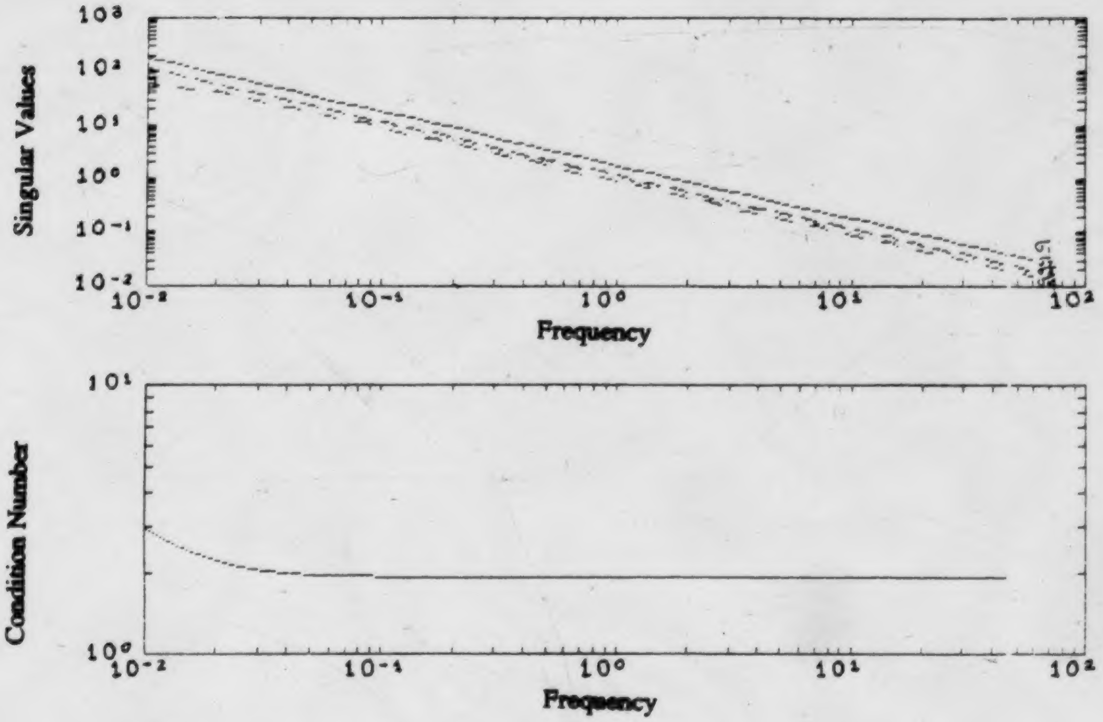

Figure 7.2: Condition Number and Singular Values of the Steady State Gain Matrix Versus Frequency for the Scaled CSTR System (Bonvin, 1985)

functions of frequency for the scaled process models. Bonvin's scaling method significantly decreased the condition number of the process model at steady state. At high frequencies, the system has a low condition number but the singular values have also decreased significantly. Therefore, this system should be able to handle high frequency disturbances easily.

7.5.2 Physical Scaling Method Based on the Relative Significance of each Variable

Johnston (1987) and Wolfgang (1986) proposed a scaling procedure based on physical considerations where the relative importance of changes in the variables are accounted for. This scaling method incorporates information relating to the importance of each input and output variable and the size of disturbances which a given control system can handle. Therefore, this scaling method can be very subjective depending on the extent of knowledge of the process.

The controlled variables are scaled such that a change of given magnitude will have equal significance for all the outputs from a controllability point of view. For example, a 10% change in distillation column temperature is clearly more critical than a 10% change in the level of liquid in a reflux accumulator. If a 10% change in the temperature was as critical as a 0.5m change in the liquid level, then the scaling factor to apply would be 0.005 to the temperature output (Johnston, 1987).

Manipulated variables are scaled such that a change of given magnitude in all the variables represents an equivalent amount of control action. If The manipulated variables represent valve movement, they can be represented by % changes from their steady state values or from their maximum range.

Disturbances should be scaled according to the expected magnitude and likelihood of each disturbance. If the disturbances were not scaled, the singular values of the load transfer function would be unrealistically influenced by the effects of the disturbance with the largest magnitude. If a change of 1% in one disturbance variable will be as likely to occur as a change of 1% in another disturbance variable, then the disturbance model expressed in these units will be properly scaled. However, if the units of one of the disturbance variables is altered, then a change of magnitude 1 in the new units will be more or less likely to occur than a 1% change in the other variable. In this case, the disturbance model will be improperly scaled. Alternatively, the disturbances can be scaled with respect to their expected magnitudes (weighted) such that changes of equal magnitude in all the disturbances can be expected.

Example 7.6: The disturbance variables w_1, w_2, and w_3, with expected ranges (in % of steady state) ∓20%, ∓20% and ∓3.4%, respectively, can be normalized with respect to their steady state values (Johnston, 1987). Then, the normalized variables are adjusted with scaling factors to account for the variation in the expected magnitudes.

disturbance	weighting factor
w_1	1.0
w_2	1.0
w_3	0.17 = 3.4/20

If the disturbance model was not scaled, the singular values of $\underline{G}_L$ would be unrealistically influenced by w_3. Therefore, the effect of the disturbance on the output would be

$$y_1(s) = g_{13}(s)w_3(s) \qquad [7.5.20]$$

With scaling,

$$w_3 = 0.17\ w_3 \qquad [7.5.21]$$

and

$$y_1(s) = 0.17g_{13}(s)w_3(s) \qquad [7.5.22]$$

Therefore, a 3.4% change in w_3 is "equivalent" to a 20% change in w_1 and w_2 and $\bar{w}_3$.

7.6 Summary

All the analysis methods based on singular value decomposition are dependent on the scaling of the matrix. Therefore, it is necessary that the process variables are scaled in a consistent manner. All the techniques described in this chapter scale the variables with respect to some reference value. Scaling with respect to a steady state value or a maximum value can alter the relative significance of each variable because they may not be referenced to the same basis point. Empirical scaling methods such as equilibration or geometric mean scaling reference all the controlled variables to the same basis point and all the manipulated variables to a reference point. However, the order of scaling (row then column or vice versa) may affect the resulting scaled values. Normalization followed by equilibration of the manipulated variables will reference the variables to a the same basis point and will eliminate the inconsistency in the order of scaling.

Numerical techniques which minimize a measure of variability may not relate directly to the process being described. If an improper criteria is used, the relative effects of the variables can be altered since these methods serve to minimize the variability of the entries in the matrix. The scaling method of Bonvin will not change the relative magnitudes of the variables because the variables are referenced to the same physical quantity. The scaling method of Johnston is also based on physical considerations but requires a good knowledge of the dynamics of the process. Therefore, this method of scaling is very subjective and may introduce inconsistencies depending on the degree of process knowledge available.

Chapter 8 Singular Value Decomposition in Predictive Controller Design

8.0 Introduction

Maurath (1985a) and Callaghan (1986) used principal components analysis in the design of long range predictive controllers based on dynamic matrix control theory. In dynamic matrix control (DMC), developed at Shell Research (Cutler, 1977, 1982), a prediction of future outputs over a prediction horizon is obtained from a step response convolution model of the process assuming no further control action is taken. The error between the predicted output trajectory and a reference trajectory is then minimized by varying a finite number of control actions (control horizon). SVA is used to determine the number of principal components of the dynamic matrix, formed from the step response coefficients, to retain in the approximate process inverse used in the controller. This procedure will produce a reduced rank approximation matrix to the original dynamic matrix which will improve the conditioning of the DMC algorithm.

8.1 Dynamic Matrix Control

The control law used in this set of predictive controllers is derived from the minimization of an appropriate performance index with respect to the future control actions. The performance index is a function of the predicted output errors over the prediction horizon and the present and future control actions. For an unconstrained controller, the performance index can be expressed as a generalized least squares problem. However, rigorous handling of process constraints on the controlled and/or manipulated variables within the performance index requires other solution methods such as linear or quadratic programming. For example, Shell has implemented a quadratic formulation for a predictive controller (QDMC) which uses a modified DMC performance index to maintain projections of any constrained variables within bounds (Garcia, 1986).

The control law used in unconstrained DMC is derived from the minimization of the following performance index (Cutler, 1979):

$$J(\Delta\underline{u}) = \left[\underline{y}_s - \left\{ \underline{y}_{OL} - \underline{A}\cdot\Delta\underline{u} \right\} \right]^T \underline{Q}^T \underline{Q} \left[\underline{y}_s - \left\{ \underline{y}_{OL} - \underline{A}\cdot\Delta\underline{u} \right\} \right] \qquad [8.1.1]$$

where

$\underline{y}_s$ = vector of setpoint values for the next R sampling intervals

$\underline{y}_{OL}$ = vector of predicted outputs for the next R sampling intervals assuming no future control actions (Δu) have been taken

Δu = vector of next L input changes (to be calculated)

$\underline{A}$ = RxL dynamic matrix, a lower triangular matrix of the step response coefficients, $\underline{A}\Delta\underline{u}$ is the effect of the next L input changes on the output at the next R sampling intervals

$\underline{Q}$ = diagonal, positive, definite weighting matrix

R = prediction horizon

L = control horizon

If

$$\underline{E} = \underline{y}_s - \underline{y}_{OL} \qquad [8.1.2]$$

represents the predicted deviations (or errors) of the output variable from its setpoint over the prediction horizon assuming no future control action is taken (open loop prediction), then the performance index can be written as

$$J(\Delta\underline{u}) = [\underline{Q}\cdot\underline{A}\cdot\Delta\underline{u} - \underline{Q}\cdot\underline{E}\ ]^{T}[\underline{Q}\cdot\underline{A}\cdot\Delta\underline{u} - \underline{Q}\cdot\underline{E}\ ] \qquad [8.1.3]$$

for which the minimization is a standard least squares problem. From equation 2.2.20, if $\underline{G} = \underline{Q}\cdot\underline{A}$, $\underline{x} = \Delta\underline{u}$, and $\underline{b} = \underline{Q}\cdot\underline{E}$, then the control law becomes

$$\Delta\underline{u} = [\underline{A}^{T}\underline{Q}^{T}\underline{QA}]^{-1}\underline{A}^{T}\underline{Q}^{T}\underline{QE} = (\underline{QA})^{-1}\underline{Q}\cdot\underline{E} \qquad [8.1.4]$$

This control law can be used for both SISO and MIMO systems. A multivariable dynamic matrix can be defined as (Garcia, 1986):

$$\underline{A} = \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} \cdots & \underline{A}_{1n} \\ \underline{A}_{21} & \underline{A}_{22} \cdots & \underline{A}_{2n} \\ \vdots & \vdots & \vdots \\ \underline{A}_{m1} & \underline{A}_{m2} \cdots & \underline{A}_{mn} \end{bmatrix} \qquad [8.1.5]$$

where $\underline{A}_{ij}$ is the dynamic matrix of dimension RxL relating the i^{th} output to the j^{th} output. Similarly, the corresponding vector of control moves is given by (Garcia, 1986):

$$\Delta\underline{U} = [\Delta\underline{u}_1^{T}\ \ \Delta\underline{u}_2^{T} \cdots \Delta\underline{u}_n^{T}]^{T} \qquad [8.1.6]$$

and the output projection error is given by (Garcia, 1986):

$$\underline{E} = [\underline{E}_1^{T}\ \ \underline{E}_2^{T} \cdots \underline{E}_m^{T}]^{T} \qquad [8.1.7]$$

Therefore, equation 8.1.4 can be used as a control law for a multivariable system.

If R and L become large or the size of the multivariable system becomes large, the vectors and matrices in the control law become large. The dynamic matrix $\underline{A}$ can approach singularity and the system can become ill-conditioned. In this case, the determination of $\Delta\underline{u}$ becomes very sensitive to small variations in either the control error vector or the process model, $\underline{A}$. In noisy systems, the result can be large changes in $\Delta\underline{u}$, the control input vector, which in most real systems is undesirable. In addition, the solution of $\Delta\underline{u}$ is numerically unreliable and can be corrupted.

8.2 Principal Components Analysis for DMC

Maurath (1985a) discusses two methods which can be used to avoid this problem when designing the controller. One method, ridge regression, weights the input changes in the performance index to decrease the required control actions. The performance index becomes,

$$J_{RR}(\Delta\underline{u}) = J(\Delta\underline{u}) + f\Delta\underline{u}^T\Delta\underline{u} \qquad [8.2.1]$$

where f is the "ridge regression parameter" or weighting factor. The resulting control law is,

$$\Delta\underline{u} = (\underline{A}^T\underline{Q}^T\underline{Q}\underline{A} - f\underline{I})^{-1}\underline{A}^T\underline{Q}^T\underline{Q}\underline{E} \qquad [8.2.2]$$

This method adds another tuning knob, f, to the control algorithm.

Alternatively, principal component analysis can be used to determine the number of principal components of the systems generalized inverse

$$(\underline{Q}\underline{A})^{-I} = [\underline{A}^T\underline{Q}^T\underline{Q}\underline{A}]^{-1}\underline{A}^T\underline{Q}^T \qquad [8.2.3]$$

to retain in the process inverse used in the controller to avoid an ill-conditioned system (Maurath, 1985a). In this method, a lower rank approximation to the original dynamic matrix $\underline{A}$ is obtained by retaining only the principal components which will balance robustness and performance characteristics of the controller. If

$$\underline{Q}\cdot\underline{A} = \underline{Z}\cdot\underline{\Sigma}\cdot\underline{V}^T \qquad [8.2.4]$$

$$\underline{w} = \underline{V}^T\Delta\underline{u} \qquad [8.2.5]$$

and

$$\underline{g} = \underline{Z}^T \underline{Q} \cdot \underline{E} \qquad [8.2.6]$$

the performance index given in equation 8.1.3 becomes

$$J(\underline{w}) = \Sigma\underline{w} - \underline{g}\;{}_2^2 \qquad [8.2.7]$$

The minimization of $J(\underline{w})$ is a least squares problem. The solution to equation 8.2.7 is

$$\underline{w} = \underline{\Sigma}^{-1}\underline{g} = \underline{V}^T \Delta\underline{u} \qquad [8.2.8]$$

In the control application, the error vector, $\underline{E}$, changes at each sampling interval and is included in the data set. However, the component selection procedure requires that the error vector $\underline{E}$ be specified so that the minimization in equation 8.2.8 can be calculated. In most cases, $\underline{E}$ will be specified so that the controller is optimized for setpoint changes.For example, in the 2x2 case,

$$\underline{E}_1 = [1\ 1\ 1\ \ldots\ 1\ 0\ 0\ \ldots\ 0]^T \qquad [8.2.9a]$$

and

$$\underline{E}_2 = [0\ 0\ 0\ \ldots.\ 0\ 1\ 1\ \ldots.\ 1] \qquad [8.2.9b]$$

From equation 8.2.8, it can be seen that $\underline{w}$ is a linear combination of the columns of $\underline{V}$ and the control actions $\Delta\underline{u}$. Each component in $\Delta\underline{u}$, Δu_i, has been projected onto the orthonormal basis set $\underline{V} = [\underline{v}_1^T\ \underline{v}_2^T\ \cdots\ \underline{v}_n^T]$ into orthogonal components w_i,referred to as principal components, with a magnitude of σ_i. Each component, w_i, contributes to improving the solution, $\Delta\underline{u}$, and the resulting control performance by decreasing the residual in the performance index (Maurath, 1985). If a component is included in the solution rather than setting its magnitude to zero, the objective function in equation 8.2.7 will be improved by g_i^2. Therefore, a convenient measure of the effect of a component on the control performance is (Maurath, 1985):

$$\left(G_F\right)_i = \frac{g_i^2}{g^T g} \qquad [8.2.10]$$

This measure has been normalized by the largest possible residual which occurs if all components are included in the solution. If the current and all previous components are retained in the solution, the remaining residual can be determined from (Maurath, 1985):

$$\left(G_S\right)_i = 1 - \frac{\sum_{j=1}^{i} g_j^2}{g^T g} \qquad [8.2.11]$$

Each component w_i contributes to the norm of the solution vector which increases the size of the control action $\Delta \underline{u}$. The first control move, which is usually the largest move, can be determined from (Maurath, 1985):

$$\left(\Delta u_1\right)_i = \sum_{j=1}^{i} v_{j1} \cdot \frac{g_j}{\sigma_j} \qquad [8.2.12]$$

Therefore, equation 8.2.12 will indicate the control energy required. In addition, each component has an associated magnitude which serves to increase the condition number of $\underline{Q} \cdot \underline{A}$. As the condition number of the dynamic matrix increases, the sensitivity of the resulting controller to modeling errors and noise increases resulting in a decrease in the robustness of the controller. Therefore, as the number of principal components included in the controller increases, there is a tradeoff between control system performance and robustness.

Removal of principal components corresponding to relatively small component magnitudes will slightly increase the residual but will significantly decrease the condition number of the dynamic matrix, $\underline{A}$ and $\underline{QA}$.

In addition, the required control action will be decreased. If k components are retained in $\underline{QA}$, the control law in equation 8.2.6 becomes,

$$(\Delta \underline{u})_k = \underline{V} \cdot \underline{\Sigma}_k^{-1} \underline{Z}^T \underline{Q} \cdot \underline{E} \qquad [8.2.13a]$$

where

$$\left(\underline{\Sigma}_k^{-1}\right)_{ii} = \begin{cases} 1/\sigma_i & i \geq k \\ 0 & i < k \end{cases} \qquad [8.2.13b]$$

and

$$\left(\underline{\Sigma}_k^{-1} \right)_{ij} = 0 \qquad i \neq j \qquad [8.2.13c]$$

Elimination of principal components in the controller design does not correspond directly to eliminating the corresponding step response coefficients in the process model. As the number of step response coefficients increases, more of the time domain information will be incorporated into the control law resulting in changes to the magnitudes of the components. When the number of step response parameters used in the dynamic matrix is sufficient to capture all the process information, the magnitude of the principal components will not change if the sequence length is increased. As the control horizon, L, is decreased, the condition number of the system decreases because the number of principal components (p*L where p is the number of manipulated variables) decreases. Therefore, decreasing the control horizon is equivalent to decreasing the number of principal components used in the DMC controller. Also, decreasing the control horizon will change the magnitude of the components.

Example 8.1: The transfer function matrix for a methanol-water distillation column is given by (Wood, 1973):

$$\begin{bmatrix} x_D \\ x_B \end{bmatrix} = \begin{bmatrix} \dfrac{12.8e^{-s}}{16.7s + 1} & \dfrac{-18.9e^{-3s}}{21s + 1} \\ \dfrac{6.6e^{-7s}}{10.9s + 1} & \dfrac{-19.4e^{-3s}}{14.4s + 1} \end{bmatrix} \cdot \begin{bmatrix} R_F \\ V \end{bmatrix} \qquad [8.2.14]$$

where x_D and x_B are the distillate and bottoms compositions, respectively, and R_F and V are the reflux and boilup rate, respectively. For a sampling time of 2.5 minutes, the principal components were determined for the system with 60 terms in the step response models and

$$R = 25$$

$$L = 10$$

Equal output weighting was used ($\underline{Q} = \underline{I}$). Table 8.1 and Table 8.2 show the magnitudes of the principal components and their corresponding contributions to the residual, the control input vector and the condition number of the dynamic matrix (equations 8.2.10, 8.2.11 and 8.2.12) for a step change in $\underline{R}_F$ and V, respectively. The 7th through 20th principal components contribute very little to the control performance as indicated by G_F and G_S but significantly increase the condition number and the control energy required ($\Delta\underline{u}_1$). Therefore, a DMC controller designed using 6 principal components will exhibit relatively good performance with good robustness.

8.3 Summary

Principal components analysis can be used to increase the robustness of DMC by decreasing the condition number of the dynamic matrix and minimizing the control action. The controller will become less sensitive to modeling errors and noise in the system. However, the control performance will be sacrificed. If the principal components which are neglected from the controller design have relatively small magnitudes, the decrease in control performance will be very small relative to the increase in robustness.

Elimination of principal components does not correspond directly to reducing the number of step response coefficients used in the process models. It removes the portion of the input and output subspaces which contribute little to the behaviour of the process. Reduction in the size of the control horizons will also eliminate principal components from the DMC controller.

Table 8.1: Principal Component Magnitudes, $(G_F)_i$, $(G_S)_i$, $\Delta\underline{u}_1$, and Condition Number of the Dynamic Matrix for a Step Change in R for Example 8.1

i	σ_i	G_F	G_S	$(\Delta R_F)_1$	$(\Delta V)_1$	$\gamma(Q\underline{\Delta})$
1	305.5	0.374	0.626	0.002	-0.003	1.0
2	51.9	0.429	0.197	0.026	-0.012	5.9
3	47.9	0.047	0.151	0.027	-0.023	6.4
4	14.4	0.022	0.129	0.037	-0.024	21.2
5	9.3	0.066	0.062	0.094	0.001	32.8
6	6.0	0.004	0.058	0.098	-0.021	50.9
7	4.0	0.020	0.038	0.169	-0.012	76.4
8	3.1	0.000	0.038	0.168	-0.016	98.5
9	2.5	0.010	0.028	0.237	-0.011	122.2
⋮	⋮	⋮	⋮	⋮	⋮	⋮
20	0.2	0.000	0.001	0.875	0.006	1528.0

Table 8.2: Principal Component Magnitudes, $(G_F)_i$, $(G_S)_i$, $\Delta\underline{u}_1$, and Condition Number of the Dynamic Matrix for a Step Change in V for Example 8.1

i	σ_i	G_F	G_S	$(\Delta R_F)_1$	$(\Delta V)_1$	$\gamma(Q\underline{\Delta})$
1	305.5	0.365	0.635	0.002	-0.003	1.0
2	51.9	0.107	0.528	-0.010	-0.001	5.9
3	47.9	0.381	0.147	-0.012	-0.026	6.4
4	14.4	0.023	0.124	-0.001	-0.049	21.2
5	9.3	0.035	0.089	-0.042	-0.067	32.8
6	6.0	0.018	0.070	-0.034	-0.113	50.9
7	4.0	0.003	0.068	-0.060	-0.116	76.4
8	3.1	0.012	0.056	-0.080	-0.183	98.5
9	2.5	0.000	0.056	-0.085	-0.183	122.2
⋮	⋮	⋮	⋮	⋮	⋮	⋮
20	0.2	0.001	0.040	-0.024	-0.390	1528.0

Chapter 9 Conclusions and Recommendations for Further Work

9.1 Conclusions

Singular value decomposition is a robust numerical tool which can be used in the analysis and design of process control systems. The analysis techniques provide a geometric interpretation of the subspaces associated with a process. However, the results from any SVD analysis can be modified by altering the scaling of the process variables. This thesis is a critical review of these SVD techniques. All the areas in which SVD techniques can be applied, with the exception of stability and robust analysis, have been discussed.

The main conclusions of this thesis are:

1. The model identification techniques using singular value decomposition of the impulse hankel matrix, with the exception of Kung's method (1978), provide a numerically robust method of identifying a process model from an impulse response. The impulse response can be obtained directly from the process step or pulse response or by time series analysis of a PRBS input. The effects of noise on the system can be filtered out and an approximate process model can be obtained. An estimate of the approximation error for these models can be calculated. Therefore, the potential robustness of a controller designed with these models is known as well as an indication of performance. Since the identified models are approximately balanced (except Kung's), they produce good scaling of the state space.

These models will predict the process response several steps into the future and therefore can be used for one-step ahead or long range predictive controllers such as GPC or DMC. However, if the model is of a lower order than the process, the identified models will exhibit steady state offset. As the deviation between the model order and the actual process order increases, the offset also increases. Therefore, any controller utilizing these models would have to include integral action. These techniques can be easily extended to multivariable systems and to time variant systems.

2. SVD of the dynamic matrix allows the control engineer to design a DMC controller which will be numerically robust. The sensitivity of the control to noise in the data and model errors can be reduced for large dimensional (or ill-conditioned) systems. Measures which are easy to calculate indicate the tradeoff between controller performance and robustness.

3. SVD provides a geometric interpretation of the interactions between the controlled variables and the manipulated variables in a process. A measure of the open loop interactions and approximate closed loop interactions has been developed which is equivalent to the direct Nyquist array. From the analysis, optimum variable pairings for a square or non-square multiloop control strategy are obtained.

4. Singular value methods allow one to analyze the controllability and observability of a system. These methods provide a geometric insight into controllability and observability. Balanced state space realizations provide optimal scaling of the controllable and observable subspaces.

Therefore, state space controllers designed using a balanced state space realization will be robust with respect to observability and controllability.

9.2 Recommendations for Further Work

The use of SVD in the analysis of control systems is a relatively new area of research. Some recommendations for further research include:

1. applying the identification and the interaction analysis techniques to data obtained from a real plant process.

2. determining an optimal method of scaling process variables because all the results from singular value methods are dependent on the scaling method employed.

REFERENCES

Adamjan, V.M., Arov, D.Z., Krein, M.G., "Analytic Properties of Schmidt Parts for a Hankel Operator and the Generalized Schur-Takari Problem", Math. USSR Sbornik, 15, pp. 31-73, (1971).

Al-Saggaf, U.B., Franklin, G.F., "Reduced Order Controller Design for Discrete Time Systems", IFAC 10th World Congress on Automatic Control, 8, pp. 322-327, (1987).

Anderson, ., Time Series Analysis and Forecasting: The Box-Jenkins Approach, Butterworth and Co. (Publishers) Ltd., London, Great Britain, (1976).

Arkun, Y., Ramakrishnan, S., "Structural Sensitivity Analysis in the Synthesis of Process Control Systems", Chem. Eng. Science, 39 (7/8), pp. 1167-1179, (1984).

Autonne, L.; Bull. Soc. Math. France, 30, pp. 121-133, 1902.

Barnett, S., Matrices in Control Theory, Van Nostrand Reinhold Co., London, England, pp. 130- 136, (1971).

Bequette, B.W., Edgar, T.F., "The Equivalence of Non-Interacting Control System Design Methods in Distillation", Proc. of ACC, pp.31-38, (1986).

Bequette, B.W., Horton, R.R., Edgar, T.F., "Resilient and Robust Control of an Energy Integrated Distillation Column", Proc. of ACC, pp. 1027-1033, (1987).

Bettayeb, M., Silverman, L.M., Safonov, M.G., "Optimal Approximation of Continuous Time Systems", Proc. IEEE Conf. Decision and Control, 1, pp. 195-198, (1980).

Bonvin, D., Mellichamp, D.A., "Rescaling State and Input Variables for the Purpose of Analyzing Process Models", Proc. of ACC, pp.626-631, (1985).

Bonvin, D., Mellichamp, D.A., "A Scaling Procedure for the Structural and Interaction Analysis of Dynamic Models", AIChE J., 33(2), pp. 250-257, (1987).

Bosley, M.J., Lees, F.P., "A Survey of Simple Transfer Function Derivations from High-Order State Variable Models", Automatica, 8, pp. 765-775, (1972).

Boyce, W.E., DiPrima, R.C., "Elementary Differential Equations and Boundary Value Problems", 3rd Edition, John Wiley and Sons, USA, pp. 292-297, (1977).

Bristol, E.H., "On a New Measure of Interaction for Multivariable Process Control", IEEE Trans. Auto. Control, AC-11, pp. 133-134, (1966).

Bristol, E.H., "RGA: Dynamic Effects of Interaction", Proc.16th IEEE Conf.Decision and Control, New Orleans, Louisiana, USA, pp. 1096-1100, (1977).

Browne, E.T., "The Characteristic Roots of a Matrix", Bull. Amer. Math. Soc., 36, pp. 707, (1930).

Bryant, G.F., Yeung, L.F., "Methods and Concepts of Dominance Optimisation", IEE Proc., Part D, 130(2), pp. 72-82, (1983).

Callaghan, P.J., Lee, P.L., "Multivariable Predictive Control of a Grinding Circuit", 3rd Conf.on Control Engineering, Inst. of Eng. Australia, Sydney, Australia, pp. 99-104, (1986).

Callaghan, P.J., Lee, P.L., "An Experimental Investigation of Predictive Controller Design by Principal Component Analysis", Dept. of Chem. Eng., University of Queensland, Australia, to be published.

Chang, B.C., "A Stable State Space Realization in the Formulation of H^{∞} Norm Computation", IEEE Trans. Automat. Contr., AC-32(9), pp. 811-815, (1987).

Chen, C.F., "Model Reduction of Multivariable Control Systems by Means of Matrix Continued Fractions", Int. J. Control, 20(2), pp. 225-238, (1974).

Chen, C.F., Shieh, L.S., "A Novel Approach to Linear Model Simplifications", Int. J. Control, 8, pp. 561, (1968).

Clarke, D.W., Mohtadi, C., Tuffs, P.S., "Generalized Predictive Control - Part I: The Basic Algorithm", Automatica, 23(2), pp. 137-148, (1987).

Clarke, D.W., Mohtadi, C., Tuffs, P.S., "Generalized Predictive Control - Part II: Extensions and Interpretations", Automatica, 23(2), pp. 149-160, (1987).

Cutler, C.R., Ramaker, B.L., "Dynamic Matrix Control - A Computer Control Algorithm", Paper Presented at Houston AIChE Meeting (86th), (1979).

Cutler, C.R., "Dynamic Matrix Control of Imbalanced Systems", ISA Transactions, 21(1), pp. 1-6, (1982).

Damen, A.A.H., Hajdasinski, A.K., "Practical Tests with Different Approximate Realizations Based on the Singular Value Decomposition of the Hankel Matrix", IFAC Identification and System Parameter Estimation Conf., Washington, D.C., 2, pp. 903-908, (1982).

Davison, E.J., "A Method of Simplyfing Linear Dynamic Systems", IEEE Trans. Automat. Contr., AC-11, pp. 93, (1966).

Davison, E.J., "A Non-Minimum Phase Index and It's Application to Interacting Multivariable Control Systems", Automatica, 5, pp. 791-799, (1969).

Davison, E.J., "Interaction Index for Multivariable Control Systems", Proc. IEE, 117(2), pp. 459-462, (1970).

Downs, J.J., "Steady State Gain Analysis for Azeotropic Distillation", JACC, 1, paper WP-7C, (1981).

Doyle, J.C., "Achievable Performance in Multivariable Feedback Systems", IEEE Decision and Control Conf., 1, pp. 250-251, (1979).

Doyle, J.C., Stein, G., "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis", IEEE Trans. Automat. Contr., AC-26(1), pp. 4-16, (1981).

Eckart, C., Young, G., "A Principal Axis Transformation for Non-Hermitian Matrices", Bull. Amer. Math. Soc., 45, pp. 118-121, (1939).

Economou, C.G., Morari, M., "IMC 6: Multiloop Design", Ind. Eng. Chem. Process Des. Dev., 25(2), pp. 411-419, (1986).

Eydgahi, A.M., Singh, H. "Family of Reduced Order Models for Linear Multivariable Systems", Electronic Letters, 23(22), pp. 1178-1180, (1987).

Forsythe, G.E., Malcolm, M.A., Moler, C.B., "Computer Methods for Mathematical Computations", Prentice Hall Series in Automatic Computation, Englewood Cliffs, New Jersey, (1977).

Gagnepain, J.D., Seborg, D.E., "Analysis of Process Interactions with Applications to Multiloop Control Systems Design", Ind. Eng. Chem. Process Des. Dev., 21,pp. 5-11, (1982).

Garcia, C.E., Morari, M., "Internal Model Control 1:", Ind. Eng. Chem. Process Des. Dev., 21, pp.308-323, (1982).

Garcia, C.E., Morari, M., "Internal Model Control 2:", Ind. Eng. Chem. Process Des. Dev., 24, pp.472-483, (1985a).

Garcia, C.E., Morari, M., "Internal Model Control 3:", Ind. Eng. Chem. Process Des. Dev., 24, pp.484-494, (1985b).

Garcia, C.E., Morshedi, A.M., "Quadratic Programming Solution of Dynamic Matrix Control (QDMC)", Chem. Eng. Commun., 46, pp. 73-87, (1986).

Gerstle, J.G., Keeton, J.M., Moore, C.F., Bruns, D.D., "Singular Value Decomposition Controller Design Using a Reduced Order Model", Proc. ACC, 2, pp.925-930, (1984).

Gill, P.E., Murray, W., Wright, M.H., Process Optimization, Academic Press Inc., New York, NY, (1981).

Glover, K., "All Optimal Hankel-Norm Approximations of Linear Multivariable Systems and their $L\infty$ Error Bounds", Int. J. Control, 39(6), pp. 1115 - 1193, (1984).

Glover, K., "Model Reduction: A Tutorial on Hankel Norm Methods and Lower Bounds on L^2 Errors", IFAC 10th World Congress on Automatic Control, 10, pp. 288-293, (1987).

Golub, G.H., Reinsch, C., "Singular Value Decomposition and Least Squares Solutions", Numer. Math., 14, pp. 403-420, (1970).

Ho, B.L., Kalman, R.E., "Effective Construction of Linear State Variable Models from Input/Output Functions", Regelungstech, 14, pp. 545-548, (1966).

Isermann, R., Digital Control Systems, Springer-Verlag, Berlin, Heidelberg, (1981).

Jensen, N., "Nyquist Design of Multivariable Systems", Ph.D. Dissertation, University of Alberta, Edmonton, Alberta, Canada, (1981).

Jensen, N., Fisher, D.G., Shah, S.L., "Interaction Analysis in Multivariable Control Systems", AIChE J., 32(6), pp. 959-970, (1986).

Johnston, R.D., Barton, G.W., "Improved Process Conditionning Using Internal Control Loops", Int. J. Control, 40(6), pp. 1051-1063, (1984).

Johnston, R.D., Barton, G.W., "Structural Interaction Analysis", Int. J. Control, 41(4), pp. 1005-1013, (1985).

Johnston, R.D., Barton, G.W., "Design and Performance Assessment of Control Systems Using Singular Value Analysis", Ind. Eng. Chem. Res., 26(4), pp. 830-839, (1987).

Jonckheere, E., Li, R., "L^∞ Error Bound for the Phase Matching Approximation (the One Step at a Time Hankel Norm Model Reduction Version)", Int. J. Control, 46(4), pp. 1343-1354, (1987).

Kailath, T., Linear Systems, Prentice Hall Inc., Englewood Cliffs, N.J., (1980).

Keller, J.P., Bonvin, D., "Selection of Inputs for the Purpose of Model Reduction and Controller Design", IFAC 10th World Congress on Automatic Control, 8, pp. 226-231, (1987).

Klema, V.C., Laub, A.J., "The Singular Value Decomposition: Its Computation and Some Applications", IEEE Trans. Automat. Contr., AC-25(2), pp.164-176, (1980).

Kung, S., "A New Identification and Model Reduction Algorithmn via Singular Value Decomposition", Proc. Asilomar Conf. on Circuits, Systems, and Computers, Pacific Grove, Cal., pp. 705-714, (1978).

Kung, S.Y., "Optimal Hankel-Norm Model Reduction - Scaler Systems", Proc. JACC, II, paper FA8.D, (1980).

Kung, S.Y., Lin, D.W., "Optimal Hankel-Norm Model Reductions: Multivariable Systems", IEEE Trans. Auto. Control, AC-26(4), pp. 832-852, (1981).

Langman, J., "Effect of Q-Weighting on Self-Tuning Control of a Binary Distillation Column", M.Sc. Dissertation, University of Alberta, Edmonton, Alberta, Canada, (1987).

Lau, H., Alverez, J., Jensen, K.F., "Synthesis of Control Structures by Singular Value Analysis: Dynamic Measures of Sensitivity and Interaction", AIChE J., 31(3), pp.427-439, (1985a).

Lau, H., Jensen, H.K., "Evaluation of Changeover Control Policies by Singular Value Analysis: Effects of Scaling", AIChE J, 31(1), pp. 135-146, (1985b).

Laub, A.J., "Computation of Balancing Transformations", Proc. JACC., paper FA8-E, (1980).

Levien, K.L., Morari, M., "Internal Model Control of Coupled Distillation Columns", Proc. ACC, 2, pp. 1066-1071, (1985).

Li, R., Jonckheere, E., "An L^{∞} Error Bound for the Phase Approximation Problem", IEEE Trans. Auto. Control, AC-32(6), pp. 517-518, (1987).

Lieslehto, J., Koivo, H.N., "An Expert System for Interaction Analysis of Multivariable Systems", Proc. JACC, pp. 961-962, (1987).

Litz, L., Roth, H., "State Decomposition for Singular Perturbation Order Reduction - A Modal Approach", Int. J. Control, 34(5), pp. 937-954, (1981).

MacFarlane, A.G.J., Dynamical System Models, George G. Harrop and Co. Ltd., Great Britain, (1970).

MacFarlane, A.G.J., "A Frequency Response Approach to the Design of Multivariable Control Systems", AIChE Symposium Series: Chemical Process Control, 72(159), pp. 126-143, (1976).

MacFarlane, A.G.J., Kouvaritakis, B., "A Design Technique for Linear Multivariable Feedback Systems", Int. J. Control, 25(6), pp. 837-874, (1977).

MacFarlane, A.G.J., Scott-Jones, D.F.A., "Vector Gain", Int. J. Control, 29(1), pp. 65-91, (1979).

Maciejowski, J.M., Vines, D., "Decoupled Control of a Macroeconomic Model Using Frequency Domain Methods", J. Economic Dynamics and Control, 7, pp. 55-77, (1984).

Maciejowski, J.M., "Balanced Realizations in System Identification", Proc. IFAC Identification and System Parameter Estimation Conf., York, UK, 2, pp. 1823-1827, (1985).

Maurath, P.R., "Predictive Controller Design with Applications to Two-point Composition Control of Distillation Columns", Ph.D. Dissertation, University of California, Santa Barbara, (1985a).

Maurath, P.R., Mellichamp, D.A., Seborg, D.E., "Predictive Controller Design for SISO Systems", Proc. of ACC, 3, pp. 1546-1552, (1985b).

Maurath, P.R., Mellichamp, D.A., Seborg, D.E., "Predictive Controller Design by Principal Components Analysis", Proc. of ACC, 3, pp.1059-1065, (1985c).

McAvoy, T.J., Interaction Analysis, ISA Monograph Series 6, Research Triangle Park, NC, (1983).

Mohtadi-Haghighi, C., "Studies in Advanced Self-Tuning Algorithmns", Ph.D. Dissertation, New College, University of Oxford, England, (1986).

Moler, C., Little, J., Bangert, S., Kleiman, S.,PC-Matlab for MS-DOS Personal Computers, The Mathworks Inc., (1986).

Moore, B.C., "Singular Value Decomposition of Linear Systems", Proc. IEEE Decision and Control Conf., pp. 66-73, (1979a).

Moore, B.C., "Principal Component Analysis in Nonlinear Systems: Preliminary esults", Proc. IEEE Decision and Control Conf., 2, pp. 1057-1060, (1979b).

Moore, B.C., "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction", IEEE Trans. Automat. Contr., AC-26(1), pp. 17-32, (1981).

Moore, C., "Application of Singular Value Decomposition to the Design, Analysis, and Control of Industrial Processes", Proc. of ACC, pp. 643-650, (1986).

Morari, M., "Design of Resilient Processing Plants III: A General Framework for the Assessment of Dynamic Resilience", Chem. Eng. Sci., 38(11), pp. 1881-1891, (1983).

Morari, M., Grimm, W., Oglesby, M.J., Prossier, I.D., "Design of Resilient Processing Plants VII: Design of Energy Management System for Unstable Reactors - New Insights", Chem. Eng. Sci., 40(2), pp. 187-198, (1985).

Mullis, C.T., Roberts, R.A., "Synthesis of Minimum Roundoff Noise Fixed Point Digital Filters", IEEE Trans. Circ. and Systems, CAS-23(9), pp. 551-562, (1976).

Navratril, J., Lim, J., Fisher, D.G., "Disturbance Feedback in Model Predictive Control Systems", to be published.

Niederlinski, A., "A New Look at Least Squares Dynamic System Identification", Int. J. Control, 40(3), pp. 467-478, (1984).

Palazoglu, A.N., Manousiouthakis, B., Arkun, Y., "Design of Chemical Plants with Improved Dynamic Operability in an Environment of Uncertainty", Ind. Eng. Chem. Proc. Des. Dev. 24, pp. 802-813, (1985).

Pernebo, L, Silverman, L.M., "Model Reduction via Balanced State Space Representation", IEEE Trans. Auto. Control, AC-27(2), pp. 382-387, (1982).

Rijnsdorp, J.E., "Interaction in Two-Variable Control Systems for Distillation Columns I", Automatica, 1, pp. 15, (1965).

Rivera, D.E., Morari, M., Skogestad, S., "Internal Model Control 4: PID Controller Design", Ind. Eng. Chem. Process Des. Dev., 25, pp. 252-265, (1986).

Shah, S.L., Cluett, W.R., "RLS Based Estimation Schemes for Self-Tuning Control", to be published, (1988).

Shokoohi, S., Silverman, L.M., VanDooren, P.M., "Linear Time-Varying Systems: Balancing and Model Reduction", IEEE Trans. Auto. Control, AC-28(8), pp. 810-822, (1983).

Shokoohi, S., Silverman, L.M., VanDooren, P., "Linear Time-Variable Systems: Stability of Reduced Models", Automatica, 20(1), pp. 59-67, (1984).

Shokoohi, S., Silverman, L.M., "Identification and Model Reduction of Time-Varying Discrete Time Systems", Automatica, 23(4), pp. 509-521, (1987a).

Shokoohi, S., "Block Balancing of Linear Systems", J. of Franklin Inst., 324(3), pp. 453-464, (1987b).

Silverman, L.M., Bettayeb, L.M., "Optimal Approximation of Linear Systems", Proc. JACC, paper FA8-A, (1980).

Skogestad, S., "Studies on Robust Control of Distillation Column", Ph.D. Thesis Dissertation, California Institute of Technology, Pasedena, California, (1987).

Skogestad, S., Morari, M., "Control of High-Purity Distillation Columns from a Robustness Viewpoint", Proc. of ACC, pp. 1785-1790, (1987).

Sylvester, Messenger of Mathematics, 19, (1889).

Tether, A.J., "Construction of Minimal Linear State Variable Models from Finite Input-Output Data", IEEE Trans. Automat. Contr., pp. 427-436, (1970).

Tewari, S.S., Sahay, B., Prasad, T., "Model Reduction Technique Applied to a Nuclear Reactor Turbine System", Int. J. Systems Sci., 18(9), pp. 1767-1775, (1987).

Tombs, M.S., Postlethwaite, J., "Truncated Balanced Realizations of a Stable Non-Minimal State Space System", Int. J. Control, 46(4), pp. 1319-1330, (1987).

Tung, L.S., Edgar, T.F., "Dynamic Interaction Analysis and It's Applications to Distillation Column Control", 16th Conf. IEEE Decision and Control, New Orleans, pp. 107-112, (1977).

Tung, L.S., Edgar, T.F., "Analysis of Control-Output Interactions in Dynamic Systems", AIChE J., 27(4), pp. 690-693, (1981).

Vanderwaale, J., Staar, J., DeMoor, B., Lauwers, J., "An Adaptive Singular Value Decomposition Algorithmn and it's Application to Adaptive Realization", Proc. 6th Int. Conf. on Analysis and Optimization of Systems, Pt. II, Nice, France, pp. 32-47, (1984).

Verhaegen, M.H., Vandewalle, J., Bril, F., Cumps, M., "A Robust and Efficient Least Squares Estimation Algorithmn", IFAC Identification and System Parameter Estimation, York, UK, 1, pp. 861-866, (1985).

Verriest, E.I., Kailath, T., "On Generalized Balanced Realizations", IEEE Trans. Auto. Control, AC-28(8), pp. 833-844, (1983).

Wahlberg, B., "On Model Reduction in System Identification", Proc. ACC, pp. 1260-1266, (1986).

Wilkinson, J.H., The Algebraic Eigenvalue Problem, Oxford University Press, London, England, (1965).

Witcher, M.F., M^{c}Avoy, T.J., "Interacting Control Systems: Steady State and Dynamic Measurement of Interaction", ISA Trans., 16(3), pp. 35-41, (1977).

Wolfgang, M.G., Lee, P.L., Callaghan, P.J., "Practical Robust Predictive Control of a Heat Exchanger Network", Dept. of Chem. Eng., University of Queensland, Australia, to be published, (1986).

Wood, R.K., Berry, M.W., "Terminal Composition Control of a Binary Distillation Column", Chem. Eng. Sci., 28, pp. 1707-1717, (1973).

Zeiger, H.P., McEwen, A.J., "Approximate Linear Realizations of Given Dimension Via Ho's Algorithmn", IEEE Trans. Auto. Control, AC-19, pp. 153, (1974).

Appendix A Two Interacting Tank System

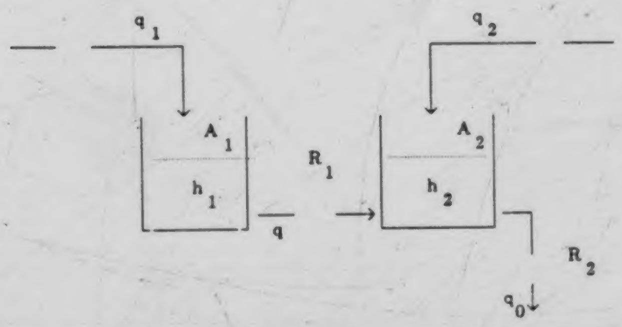

Figure A.1: Two Interacting Tanks in Series

Figure A.1 shows a system with two interacting tanks in series. Assuming linear valves in the system, the mass conservation equations for each tank are:

a) Tank 1:

$$q_1 - q = A_1 \frac{dh_1}{dt} \qquad [A.1]$$

Assuming linear valves:

$$q = \frac{h_1 - h_2}{R_1} \qquad [A.2]$$

Therefore,

$$q_1 - \frac{h_1}{R_1} + \frac{h_2}{R_2} = A_1 \frac{dh_1}{dt} \qquad [A.3]$$

Taking laplace transforms of both sides of equation A.3 and expressing the variables in deviation form,

$$Q_1(s) - \frac{H_1(s)}{R_1} + \frac{H_2(s)}{R_2} = sA_1R_1H_1(s) \qquad [A.4]$$

$$R_1Q_1(s) - H_1(s) + H_2(s) = sA_1R_1H_1 \qquad [A.5]$$

$$H_1(s) = \frac{R_1Q_1(s) + H_2}{sA_1R_1 + 1} \qquad [A.6]$$

b) Tank 2:

Similiar to tank 1, the material balance is

$$q \; + q_2 - q_0 = A_2\frac{dh_2}{dt} \qquad [A.7]$$

In terms of deviation variables in the laplace domain,

$$\frac{H_1(s)}{R_1} - \frac{H_2(s)}{R_1} + Q_2(s) - \frac{H_2(s)}{R_2} = sA_2H_2(s) \qquad [A.8]$$

$$R_2H_1(s) - (R_1 + R_2)H_2(s) + R_1R_2Q_2(s) = sA_2R_1R_2H_2(s) \qquad [A.9]$$

Substituting for $H_1(s)$ from equation A.5 into equation A.9, obtain,

$$H_2(s) = \frac{R_1R_2Q_1(s) + (sA_1R_1+1)R_1R_2Q_2(s)}{(sA_1R_1+1)(sA_2R_1R_2+R_1+R_2) - R_2} \qquad [A.10]$$

Substituting for $H_2(s)$ into equation A.5, obtain

$$H_1(s) = \frac{R_1(sA_2R_1R_2+R_1+R_2)Q_1(s) + R_1R_2Q_2(s)}{(sA_1R_1+1)(sA_2R_1R_2+R_1+R_2) - R_2} \qquad [A.11]$$

The transfer function model can be written as,

$$\begin{bmatrix} H_1(s) \\ H_2(s) \end{bmatrix} = \begin{bmatrix} g_{11}(s) & g_{12}(s) \\ g_{21}(s) & g_{22}(s) \end{bmatrix} \cdot \begin{bmatrix} Q_1(s) \\ Q_2(s) \end{bmatrix} \qquad [A.12]$$

where the elements of the process TFM, $g_{ij}(s)$, are given in equations A.11 and A.10.

Appendix B Derivation of a Model for a System of Two CSTR's in Series (Bonvin, 1985, 1987)

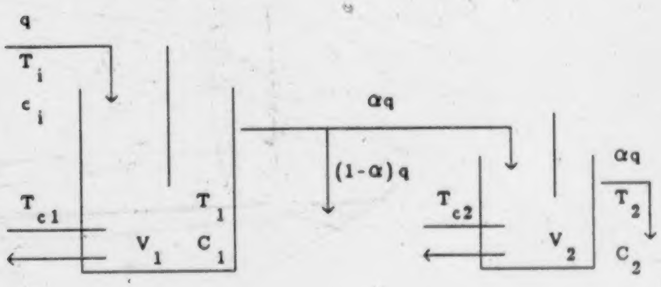

Figure B.1: Two CSTR's in Series (Bonvin, 1987)

Figure B.1 shows two continuous stirred tank heaters (CSTR) in series with a first-order exothermic reaction, jacket or coil-cooling and a by-pass around the second reactor. The conservation equations for this system are:

a) Component balances, mol/s:

$$V_1 \frac{dC_1}{dt} = q(C_i - C_1) - V_1 C_1 k_o e^{-E/RT_1} \qquad [B.1]$$

$$V_2 \frac{dC_2}{dt} = \alpha q(C_1 - C_2) - V_2 C_2 k_o e^{-E/RT_2} \qquad [B.2]$$

b) Energy balances, J/s:

$$V_1 \rho C_p \frac{dT_1}{dt} = q\rho C_p(T_i - T_1) + (-\Delta H) V_1 C_1 k_o e^{-E/RT_1} + (UA)_1 (T_{c1} - T_1) \qquad [B.3]$$

$$V_2 \rho C_p \frac{dT_2}{dt} = \alpha q\rho C_p(T_1 - T_2) + (-\Delta H) V_2 C_2 k_o e^{-E/RT_2} + (UA)_2 (T_{c2} - T_2) \qquad [B.4]$$

There are four state variables (C_1, C_2, T_2, T_1), six input variables (q, T_i, C_i, T_{c1}, T_{c2}, α) and two physical quantities (energy, number of moles of key component). It is assumed that the four state variables are measurable but the sensors give different units:

$$C_{1m}(\mathrm{mol/L}) = C_1 \qquad [B.5]$$

$$C_{2m}(\mathrm{mol/m^3}) = 1000 \times C_2 \qquad [B.6]$$

$$T_{1m}(^{o}C) = T_1 \qquad [B.7]$$

$$T_{2m}(^{o}F) = 1.8 \times T_2 \qquad [B.8]$$

Energy is chosen as the reference quantity, Π. Equations B.1 and B.2 are multiplied throughout by the equivalence factor (-ΔH). Then

$$\begin{aligned}\dot{\Pi}_1 &= (-\Delta H)V_1\dot{C}_1 = (-\Delta H)q(C_i-C_1) - (-\Delta H)V_1C_1k_oe^{-E/RT_1}\\ &= (-\Delta H)qC_i-(\Pi_1/V_1) - \Pi_1k_oe^{-E/RT_1}\end{aligned} \qquad [B.9]$$

$$\begin{aligned}\dot{\Pi}_2 &= (-\Delta H)V_2\dot{C}_2 = (-\Delta H)\alpha q(C_1-C_2) - (-\Delta H)V_2C_2k_oe^{-E/RT_2}\\ &= \alpha q(\Pi_1/V_1)-\alpha q(\Pi_2/V_2) - \Pi_2k_oe^{-E/RT_2}\end{aligned} \qquad [B.10]$$

$$\begin{aligned}\Pi_3 = V_1\rho C_pT_1 &= q\rho C_p(T_i-T_1) + (-\Delta H)V_1C_1k_oe^{-E/RT_1} + (UA)_1(T_{c1}-T_1)\\ &= q\rho C_pT_i-q(\Pi_3/V_1) + \Pi_1k_oe^{-E/RT_1} + (UA)_1(T_{c1}-\Pi_1/V_1\rho C_p)\end{aligned} \qquad [B.11]$$

$$\begin{aligned}\dot{\Pi}_4 = V_1\rho C_p\dot{T}_1 &= \alpha q\rho C_p(T_1-T_2) + (-\Delta H)V_2C_2k_oe^{-E/RT_2} + (UA)_2(T_{c2}-T_2)\\ &= \alpha q(\Pi_3/V_1-\Pi_4/V_2) + \Pi_2k_oe^{-E/RT_2} + (UA)_2(T_{c2}-\Pi_4/V_2\rho C_p)\end{aligned} \qquad [B.12]$$

where the state variables are "scaled" in terms of Π using the scaling factors

$$\Pi_1 = p_{c_1}C_1 = V_1(-\Delta H)C_1 = \tilde{x}_1 \qquad [B.13]$$

$$\Pi_2 = p_{c_2}C_2 = V_2(-\Delta H)C_2 = \tilde{x}_2 \qquad [B.14]$$

$$\Pi_3 = p_{T_1}T_1 = V_1\rho C_pT_1 = \tilde{x}_3 \qquad [B.15]$$

$$\Pi_4 = p_{T_2}T_2 = V_2\rho C_pT_2 = \tilde{x}_4 \qquad [B.16]$$

The input scaling factors, w_{si} for the s^{th} state and the i^{th} input are:

$$w_q = \max\left\{\left|(-\Delta H)(\overline{C}_i-\overline{C}_1)\right|,\ \left|(-\Delta H)\bar{\alpha}(\overline{C}_1-\overline{C}_2)\right|,\ \left|\rho C_p(\overline{T}_i-\overline{T}_1)\right|,\left|\bar{\alpha}\rho C_p(\overline{T}_1-\overline{T}_2)\right|\right\} \qquad [B.17]$$

$$w_{T_i} = \bar{q}\rho C_p \qquad [B.18]$$

$$w_{C_i} = \left|(-\Delta H)\bar{q}\right| \qquad [B.19]$$

$$w_{c_1} = (UA)_1 \qquad [B.20]$$

$$w_{c_2} = (UA)_2 \qquad [B.21]$$

$$w_{\alpha} = \max\left\{\left|(-\Delta H)\bar{q}(\overline{C}_1-\overline{C}_2)\right|,\ \left|\bar{q}\rho C_p(\overline{T}_1-\overline{T}_2)\right|\right\} \qquad [B.22]$$

The bar over the variables indicates it's steady state value.

The output scaling factors are

$$v_{c_{1m}} = \left|V_1(-\Delta H)\right| \qquad [B.23]$$

$$v_{T_{1m}} = V_1\rho C_p \qquad [B.24]$$

$$v_{C_{2m}} = \left|\frac{V_2(-\Delta H)}{1000}\right| \qquad [B.25]$$

$$v_{T_{2m}} = \frac{V_2\rho C_p}{1.8} \qquad [B.26]$$

The numerical data which Bonvin used to evaluate the models in equations B.1 to B.4 is given in Table B.1.

Table B.1: Steady State Values for the CSTR Model

Variable	Steady State (or Nominal) Values
V_1	1000 L
V_2	10L
$\bar{q}$	1 L/s
$\bar{\alpha}$	0.1
ρC_p	4000 J/L°C
$(-\Delta H)$	10^5 J/mol
$\bar{C}_i$	1 mol/L
$\bar{C}_1$	0.314 mol/L
$\bar{C}_2$	0.257 mol/L
T_i	20 °C
T_1	34°C
T_2	34.3 °C
$T_{c_1} = T_{c_2}$	10 °C
k_o	5000 s^{-1}
E/R	4500 K
$(UA)_1$	500 W/°C
$(UA)_2$	20 W/°C

Apendix C PC-Matlab Programs

```
%
 FUNCTION [r,theta] = ellipse(a,b,phi)
%
%   This program will generate an ellipse with major axis length, a
%   minor axis length, b, and rotated from the x-axis by an angle phi.
%
%     Variables:
%
%       a         = length of the major axis
%       b         = length of the minor axis
%       phi       = angle of rotation from the x-axis
%       r         = radius of the ellipse, function of theta
%       theta    = correponding angles from 0 to 2*pi
%
     k = 0:0.1:2*pi+0.1;
     [m,n] = size(k);
     for i = 1:n
       r(i) = a*b/sqrt(((b^2)*(cos(k(i))^2))+((a^2)*(sin(k(i))^2)));
       theta(i) = k(i)+phi;
     end
     end
```

```
% HOKALMAN.m
%
% MODEL IDENTIFICATION USING VARIOUS HO-KALMAN ALGORITHMNS
%
%       [A,B,C,n] = hokalman(a,b,c,d,N,m,p,r) returns a rth order state space
%                       realization for the impulse response sequence Y and the
%                     rank, n, of the Hankel matrix formed from the impulse
%                    response
%
%       [A,B,C,n,H] = hokalman(a,b,c,d,N,m,p,r) also returns the Hankel
%                   matrix H
%
%       Variables:
%
%         Y       = Impulse response sequence for the system
%        Ts      = Sampling time used
%         N       = Number of data points in the impulse response
%          r        = Order of the realization
%           n        = Rank of the Hankel matrix, states in the minimal realization
%        m       = Number of outputs
%        p       = Number of inputs
%         H       = Hankel matrix , qxl matrix
%         Hs      = Shifted hankel matrix
%         M       = Size of the square hankel matrix H
%          nm       = Spectral norm of the reduction error ||H - Hr||
%        q       = rows of the hankel matrix, M*m
%        l       = columns of the hankel matrix, M*p
%
%
%       ref:
%
%               Damen, A.A.H., Hajdasinski, A.K., "Practical Tests with Different
%                         Approximate Realizations Based on the Singular Value
%                   Decomposition of the Hankel Matrix", IFAC Identification and
%                 System Parameter Estimation Conf., Washington, D.C., pp.93-101,
%                 1982
%
%             Ho, B.L., Kalman, R.E., "Effective Construction of Linear State
%                 Variable Models from Input/Output Functions", Regelungstech,
%                 14, pp. 545-548, 1966
%
%           Kung, S., "A New Identification and Model Reduction Algorithmn
%                  via Singular Value Decompositions", 1978 Asilomar Conference
%                  on Circuits, Systems and Computers, p. 705-714, 1978
%
%           Tether, A.J., "Construction of Minimal Linear State Variable Models
%                 from Finite Input-Output Data", IEEE Trans.Automat.Contr.,
%                 pp. 427-436, 1970
% ------------------------------------------------------------------------
%
% Obtain the impulse response for each input and output (discrete system)
%
%       Y = [];
%      for i = 1:p
%           y = dimpulse(a,b,c,d,i,N+1);
%         Y = [Y y];
%     end
%       Y = Y(2:N+1,:)       % the first impulse response paramete h(0)=0
%
% Obtain the impulse response for each input and output (discrete system) for
% the case with noise in the system
%
      Y = [];
      for i = 1:p
         y=dimpulse(a,b,c,d,1,N+1);
         e=rand(N+1,1);
         e=e*0.1;
         y=y+e;
         Y = [Y y];
      end
      Y = Y(2:N+1,:);
%
% Form the (block) Hankel matrix
%
%     N = N/Ts;
      M = N/2;
      M = M-1;
      J = [];
      h = [];
      for t = 1:N
         for i = 1:m;
            for j = 1:p;
               ii = (j-1)*m+i;
               h(i,j)   =   Y(t,ii);
            end
         end
         J = [J
              h];
      end
      H = J(1:m*M,:);
      for i = 2:M
         temp = [];
         temp = J(i*m:m*(M-1+i),:);
%         for ii = i*m
%            temp = [temp
%                        J(ii,:)];
%        end
         H = [H temp];
      end
      [q,l]  =  size(H);
%
%  Obtain a minimal realization from the original Ho-Kalman algorithmn
%
%      [Q,R] = qr(H);
%      n = rank(H);
%     if n < M
%           U = [eye(n,n)  zeros(n,l-n)];
%       else
%           U = eye(q,l);
%      end
%     W = U'*U;
%
%  Shift the Hankel matrix up by (m) row
%
       temp  =  [];
       Hs  =  H(m+1:q,:);
       for i = (M+1)*m:2*M*m
           temp = [temp
                   J(i,:)];
       end
       temp
%       for i =1:M
%           temp = [temp J((M+i)*m-1:(M+i)*m,:)];
%      end
       Hs = [Hs
             temp'];
       Ep  =  [eye(p,p)
              zeros(q-p,p)];              % selects first p columns
       Em  =  [eye(m,m) zeros(m,l-m)];  % selects first m rows
%     A = U*(W*Q*Hs*R'*J)*U';
%     B = U*(W*Q*H*Ep');
%     C = (Em*H*R'*W)*U';
%
%   Obtain a rth order minimal realization using Zeigler-McEwan algorithmn
%
       [Z,S,V]=svd(H);
       Sr    =    S(1:r,1:r);
       Zr    =    Z(1:q,1:r);
       Vr    =    V(1:l,1:r);
%          Sr  =  [Sr1  zeros(r,l-r)
%                  zeros(q-r,l)];
       Hr  =  Zr*Sr*Vr';
       nm = norm(H-Hr);
       temp=[];
       for  i  =  1:r
           temp(i)     =     1/sqrt(S(i,i))
       end
       invSr  =  diag(temp);
%          invSr  =  [invSr  zeros(r,l-r)
%                            zeros(q-r,l)];
       A1   =   (invSr)*Zr'*Hs*Vr*invSr;
       B1   =  (invSr)*Zr'*H*Ep;
       C1 = Em*H*Vr*invSr;
       Ba   =   sqrt(Sr)*Vr'*Ep;
       Ca   =  Em*Zr*sqrt(Sr);
%       Hr1  =  Zr*Sr1*Vr';
%            A1a  =  inv(sqrt(Sr1))*Zr'*Hs*Vr*inv(sqrt(Sr1));
%          B1a  =  inv(sqrt(Sr1))*Zr'*H*Ep;
%         C1a  =  Em*H*Vr*inv(sqrt(Sr1));
%
%   Obtain a rth order realization using Hajdasinski and Damen algorithmn
```

```
%
       h = 0;
%
%         Shift the reduced order hankel matrix up by (m) rows
%
       temp  =  [];
       Hrs   =   Hr(m+1:q,:);
       for i = (M+1)*m:2*M*m
           temp = [temp
                   J(i,:)];
       end
       Hrs = [Hrs
              temp'];
       A2  =  invSr*Zr'*Hrs*Vr*invSr;
       B2  =  invSr*Zr'*Hr*Ep;
       C2 = Em*Hr*Vr*invSr;
       altA2  =  invSr*Zr'*Hs*Vr*invSr;
       altB2  =  invSr*Zr'*H*Ep;
       altC2 = Em*H*Vr*invSr;
%
%    Obtain a rth order realization using Kung's algorithmn
%
       sq  =  sqrt(S);
       z  =  Z*sq;
       v  =  sq*V';
%
%        Shift the matrices z and v up by m rows and add m rows of zeros to the
%         shifted matrices
%
       Q  =  zeros(m,r);
       zr    =    z(1:q,1:r);
       zup   =   zr(m+1:q,:);
       zup  = [zup
               Q];
       vr    =   v(1:r,:);
       vup   =   vr(:,m+1:q);
       vup  = [vup   Q'];
       A3 = zrup;
       Ao = vup/vr;
       B3   =   vr(:,1:p);
       C3   =   zr(1:m,:);
%
%    Obtain a rth order realization using Damen and Hajdanski modified algorithmn
%
       iter  =  0;
       num = 1;
       Hra = H;
       [z,s,v]   =   svd(Hra);
       x  =  r;
       if  iter  ==  0;
           if num < 20
               zr    =    z(:,1:r);
               vr    =    v(:,1:r);
               sr    =    s(1:r,1:r);
```

```
%                 sr = [sr  zeros(r,l-r)
%                       zeros(q-r,l)];
              Hra  =  zr*sr*vr';
              temp  =  [];
              for  w  =  1:p
                 for  k  =  1:m
                    for  i  =  1:M
                       sum  =  0;
                       for  j  =  1:i
                          sum  =  sum  +  Hra((i-j)*m+k,(j-1)*p+w);
                       end
                       h(i,k)  =  sum/i;
                    end
                    for  i  =  M+1:2*M-1
                       sum  =  0;
                       for  j  =  1:2*M-i
                          jj  =  i-M+j;
                          sum  =  sum  +  Hra((i-jj)*m+k,(jj-1)*p+w);
                       end
                       h(i,k)  =  sum/(2*M-i);
                    end
                 end
                 temp  =  [temp  h];
                 K  =  temp;
              end
              Hra  =  K(1:m*M,:);
              for  i  =  2:M
                 temp  =  [];
                 for  ii  =  i*m:(M+i-1)*m
                    temp  =  [temp
                             K(ii,:)];
                 end
                 Hra  =  [Hra  temp];
              end
              [z,s,v]  =  svd(Hra);
              x  =  rank(Hra);
              if  x  ==  r
                 iter  =  1;
              end
           end
           num = num+1;
        end
%             sr  =  [s(r,r)  zeros(r,l-r)
%                     zeros(q-r,l)]
        sr    =    s(1:r,1:r);
        zr    =    z(:,1:r);
        vr    =    v(:,1:r);
        temp  =  [];
        for  i  =  1:r
           temp(i)    =    1/sqrt(sr(i,i));
        end
        invsr  =  diag(temp);
        temp  =  [];
        Hras  =  Hra(m+1:q,:);
        for i = (M+1)*m:(2*M-1)*m
            temp = [temp
                       K(i,:)];
        end
        temp = [temp
                   J(2*M*m,:)];
        Hras = [Hras
                   temp'];
        A4   =   invsr*zr'*Hras*vr*invsr;
        B4  =  invsr*zr'*Hra*Ep;
        C4  =  Em*Hra*vr*invsr';
end
```

```
FUNCTION  [Ar,Br,Cr,k,hratio]=ibmodred(A,B,C,minor,maxor,m,n,dc)
%
% INTERNALLY-BALANCED MODEL REDUCTION:
%
%       Model Reduction scheme for a state space system description using
%       internal balancing of the system (A,B,C).  Weak subsystems are
%       eliminated to produce a reduced order model (Ar,Br,Cr)
%
%         [Ar,Br,Cr,k] = ibmodred(A,B,C,minor,maxor,m,n,dc) returns an internally
%                         balanced kth order state space model of the system
%                      (A,B,C)
%
%         [Ar,Br,Cr,k,hratio] = ibmodred(A,B,C,minor,maxor,m,n,dc) also returns the
%                         ratio of the (k+1 to n) and the first k second order
%                      modes of the system, hratio
%
%     Ref:
%
%           1) Moore, B.C., "Principal Component Analysis in Linear Systems:
%              Controllability, Observability and Model Reduction", IEEE Trans
%           AC, 26-1, Feb. 1981
%
%           2) Moore, B.C., "Singular Value Analysis of Linear Systems",
%             IEEE Decision and Control Conference, Jan. 1979, pp. 66-73
%
%
%   Variables:
%
%       A,B,C          = Original state space realization
%        Aib,Bib,Cib   = Internally balanced state space realization
%       Ar,Br,Cr       = Reduced order state space realization
%      dc              = Flag indicates type of system:
%                           dc = 1, discrete time
%                           dc = 2, continous time
%      maxor          = Maximum order of the realizations
%      minor          = Minimum order of the realizations
%       k              = Order of the realizations
%      m              = Number of inputs
%      n              = Number of outputs
%       hratio         = Ratio of the k+1 to n second order modes to the first
%                          k second order modes, an error criteria
%       Wc2            = Controllability gramian
%      Wo2            = Observability gramian
%
%
      j  =  sqrt(-1);
%
%        Obtain the internally-balanced model
%
      [Aib,Bib,Cib,Wc2,Wo2,H,T]  =  intbal(A,B,C,dc);
      [uc,sc,vc]  =  svd(Wc2);
      [uo,so,vo]  =  svd(Wo2);
%
```

```
%        Calculate the second order modes
%
     [uh,sh,vh]  =  svd(H);
     dsh  =  diag(sh);
%
%        Find a representative kth order model for (Aib,Bib,Cib)
%
     L  =  length(dsh);
     for k = minor,maxor
        x1  =  sh(1:k);
        if k==L
           x2  =  [];
        else
           x2  =  sh(k+1:L);
        end
        for i = 1,k
           X1(i)  =  x1(i)^2;
        end
        if k==L
           X2 = [];
           hratio  =  inf;
        else
           ii=L-k+1;
           for i  =  1,ii
              X2(i)  =  x2(i)^2
           end
           hratio   =   sqrt(sum(X1))/sqrt(sum(X2))
        end
        Ar  =  Aib(1:k,1:k);
        Br  =  Bib(1:k,1:m);
        Cr  =  Cib(1:n,1:k);
        if dc == 1
           c  =  ctrb(Ar,Br);
           o  =  obsv(Ar,Cr);
           Wc2 = c*c';
           Wo2 = o'*o;
        else
           Wc2 = gram(Ar,Br);
           Wo2  =  gram(Ar',Cr');
        end
        k
        Ar
        pause
        Br
        pause
        Cr
        pause
        fprintf('Controllability  gramian  of  reduced  order  system')
        Wc2
        pause
        fprintf('Observability  gramian  of  reduced  order  system')
        Wo2
        pause
        hratio
```

```
   pause
end
mo = diag(dsh);
fprintf('Diagonal matrix of the second order modes of system')
mo
end
```

```
FUNCTION  [Ar,Br,Cr,k,hratio]=inmodred(A,B,C,minor,maxor,dc)
%
% INPUT-NORMAL MODEL REDUCTION:
%
%       Model Reduction scheme for a state space system description using
%       input normalization of the system (A,B,C).  Weak subsystems are
%       eliminated to produce a reduced order model (Ar,Br,Cr)
%
%        [Ar,Br,Cr,k] = inmodred(A,B,C,minor,maxor,dc) returns a input normalized
%                       kth order state space model of the system (A,B,C)
%
%         [Ar,Br,Cr,k,hratio] = inmodred(A,B,C,minor,maxor,dc) also returns the
%                        ratio of the (k+1 to n) and the first k second order
%                       modes of the system, hratio
%
%      Ref:
%
%           1) Moore, B.C., "Principal Component Analysis in Linear Systems:
%               Controllability, Observability and Model Reduction", IEEE Trans
%            AC, 26-1, Feb. 1981
%
%           2) Moore, B.C., "Singular Value Analysis of Linear Systems",
%             IEEE Decision and Control Conference, Jan. 1979, pp. 66-73
%
%
%       Variables:
%
%        A,B,C            = Original state space realization
%          Ain,Bin,Cin  = Input normalized state space realization
%         Ar,Br,Cr       = Reduced order input normalized state space realization
%       dc              = Flag indicates type of system:
%                              dc = 1, discrete time
%                              dc = 2, continous time
%       minor           = Minimum order of realization
%       maxor           = Maximum order of realization
%        k                = Order of realization
%       hratio          = Ratio of the sum of the (k+1) to n second order modes
%                             and the first k second order modes, an error criteria
%       Wc2              = Controllability gramian
%       Wo2             = Observability gramian
%
%
        j  =  sqrt(-1);
%
%        Obtain the input-normalized model
%
        [Ain,Bin,Cin,Wc2,Wo2,H,T]  =  inpnorm(A,B,C,dc);
        [uc,sc,vc]=svd(Wc2)
%
%        Calculate the second order modes
%
        [uh,sh,vh]  =  svd(H);
        dsh  =  diag(sh);
```

```
%
%      Find a representative kth order model for (Ain,Bin,Cin)
%
     n = length(dsh);
     for k = minor,maxor
        x1 = sh(1:k);
        x2 = sh(k+1:n);
        for i = 1,k
           X1(i) = x1(i)^2;
        end
        ii = n - (k+1)
        for i = 1,ii
           X2(i) = x2(i)^2;
        end
        hratio = sqrt(sum(X1))/sqrt(sum(X2));
        Ar = uc'*Ain*uc;
        Br = uc'*Bin;
        Cr = Cin*uc;
        if dc == 1
           c = ctrb(Ar,Br);
           o = obsv(Ar,Cr);
           Wc2 = c*c';
           Wo2 = o'*o;
        else
           Wc2 = gram(Ar,Br),
           Wo2 = gram(Ar',Cr');
        end
        mo = diag(dsh)
        fprint('Observability gramian of reduced order system')
        Wo2
        fprint('Diagonal matrix of the second order modes')
        mo
        pause
     end
```

```
FUNCTION [Ain,Bin,Cin,Wc2,Wo2,H,T] = inpnorm(A,B,C,dc)
%
% INPUT-NORMAL  State-space realization with input normalized internal
%                  coordinate system
%                   ie.  Wc2 = I,  Wo2 = S4  S2=second order modes
%
%                   [Ain,Bin,Cin] = inpnorm(A,B,C,dc) returns a input normalized
%                        state-space representation of the system (A,B,C)
%
%                  [Ain,Bin,Cin,Wc2,Wo2,H,T] = inpnorm(A,B,C,dc) also returns
%                        matrices Wc2 and Wo2, the controllability and observability
%                          gramian, respectively, of the normalized realization,
%                      matrix H, the Hankel matrix and matrix T, the simularity
%                        transformation used to convert (A,B,C) to (Ain,Bin,Cin)
%
%
%               Ref:
%                    1) Moore, B.C., Principal Component Analysis in Linear
%                           Systems:  Controllability, Observability, and Model
%                     Reduction, IEEE Trans.AC, 26-1, Feb. 1981
%
%                     2) Moore, B.C., Singular Value Analysis of Linear Systems,
%                        IEEE Decision and Control Conference, Jan. 1979, pp.66-73
%
%                     3) Laub, A.J., "Computation of Balancing Transformations",
%                     Proc. JACC Conf., Vol. 1, paper FA8-E, 1980
%
%           see: balreal.m, outnorm.m, intbal.m
%
%    Variables:
%
%       A,B,C           = Original state space realization of process
%       Ain,Bin,Cin  = Input normalized realization
%      dc               = Flag which indicates the type of system:
%                              dc = 1, discrete time
%                              dc = 2, continuous time
%     Wc2              = Controllability gramian
%     Wo2             = Observability gramian
%     H               = Hankel matrix of system
%     T               = Transformation matrix
%
%
     j  =  sqrt(-1);
%
%    Calculate  the  observability  and  controllability  gramians
%
%    a) Discrete  case:
%
     if dc == 1
        c  =  ctrb(A,B);
        o  =  obsv(A,C);
        wc2  =  c*c';
        wo2  =  o'*o;
        else
            wc2 = gram(A,B);
            wo2 = gram(A',C');
        end
%
% Find the singular value decomposition of the gramians.  Note: the right and
% left singular vectors of the gramians are equal to the eigenvectors and the
% singular values of the gramians are equal to their respective eigenvalues
%
        [uc,sc,vc] = svd(wc2);
        ssc = sqrt(sc);
        [uo,so,vo] = svd(wo2);
        sso = sqrt(so);
%
% Find the Hankel matrix which describes the system and determine the
% second order modes (the singular values of H) Also, calculate the
% transformation matrix T.  Note: the second order modes are invariant under
%  internal coordinate transformations
%
        H = sso*vo'*vc*ssc;
        [uh,sh,vh] = svd(H);
        ssh = sqrt(sh);
        T = uc*ssc*vh;
%
% The input normalized state space realization
%
        Ain = T*T;
        Bin = T;
        Cin = C*T;
        if nargout > 3
            if dc == 1
                c = ctrb(Ain,Bin);
                o = obsv(Ain,Cin);
                Wc2 = c*c';
                Wo2 = o'*o;
            else
                Wc2 = gram(Ain,Bin);
                Wo2 = gram(Ain',Cin');
            end
        end
        mc = eye(Wc2);
        mo = diag(diag(sh^2));
%       if mc ~= Wc2
%               fprintf('Input Normalization Incorrect')
%           mc
%           Wc2
%       end
%       if mo ~= Wo2
%               fprintf('Input Normalization Incorrect')
%           mo
%           Wo2
%       end
        end
```

```
% INTACT.M
%
%       This program analyzes interactions in multivariable systems
%      with transfer function G(s) where s=iw using the direct nyquist
%      array (DNA) and the interaction measure of Lau (alignment angles)
%
%                 num
%         G(s) =  ---
%                 den
%
%
%      Variables:
%
%   G        = transfer function matrix
%   Z        = left singular vectors of G
%   S       = singular values of G
%   V        = right singular vectors of G
%   W        = rotational matrix
%   w      = frequency
%    theta  = alignment angles, radians
%    deg    = alignment angles, degrees
%    re      = real part of nyquist plot
%    im      = imaginary part of nyquist plot
%    i       = imaginary number, sqrt(-1)
%    s        = laplace transform variable, s = iw
%
        i  =  sqrt(-1);
%
%    Input the transfer function matrix
%
        m = input('Number of inputs, m = ')
        n = input('Number of outputs, n = ')
        fprintf('Input the elements of the transfer function matrix')
        fprintf('row by row starting with G(1,1) - G(1,m) and ending')
        fprintf('with  G(n,1)  -  G(n,m)')
        num = [];
        den = [];
        for k = 1:n
   for j = 1:m
      temp = input('num = ')
      num = [num
      temp];
      temp = input('den = ')
      den = [den
      temp];
   end
        end
%
%    Input the frequency range of interest
%
        x1 = input('minimum frequency is 10 to the power x1 = ')
        x2 = input('maximum frequency is 10 to the power x2 = ')
        w  =  logspace(x1,x2);
```

```
%
%   Calculate the alignment angles
%
      s = i*w;
      W = [];
      sigma = [];
      for k = 1:50
%
%     Calculate the polynomial TFM at s=i*w
%
  j = 1;
  for r = 1:n
     for c = 1:m
        G(r,c)     =     polyval(num(j,:),s(k))/polyval(den(j,:),s(k));
        j = j+1;
     end
  end
  [Z,S,V] = svd(G);
%
%      Calculate the radius of the Gershgorin bands at this frequency
%
  sum = 0;
  for c = 1:m
     for r = 1:n
        if r ~= c
   sum    =    sum+sqrt((real(G(r,c))^2)+(imag(G(r,c))^2));
        end
     end
     radius(k,c) = sum;
  end
  temp = diag(S);
  sigma = [sigma temp];
%
%       Calculate the rotational matrices for the n loops (l)
%
%  row = [3 2 1];
%  col = [2 1 3];
  for l = 1:n
     tempw   =   Z(:,l)*V(:,l)';
%
%      Check for the largest entry in tempw, x(i,j) corresponding to a
%      u(i) - y(j) pairing
%
     x = 0;
     for r = 1:n
        for c = 1:m
    y  =  abs(tempw(r,c));
    if y > x
      x = y;
      row(k,l) = r;
      col(k,l) = c;
    end
        end
     end
```

```
%
%      Calculate  the  alignment  angle  for  pairing  u(r)  -  y(c)
%
     theta(k,l)     =     acos(sqrt(tempw(row(k,l),col(k,l))     ...
    *conj(tempw(row(k,l),col(k,l)))));
%           theta(k,l )   =   acos(sqrt(tempw(row(l),col(l))   ...
%         *conj(temp w (row(l),col(l)))));
     deg(k,l)     =     360*theta(k,l)/(2*pi);
     W = [W tempw];
  end
%
%    Calculate  the  condition  number  of  G(iw)
%
  cond(k)    =    sigma(1,k)/sigma(n,k);
%
%     Calculate  the  total  interaction  measure  at  w
%
  sum1 = 0;
  sum2 = 0;
  for  j  =  1:n
      sum1      =      sum1+((sigma(j,k)^2)*(cos(theta(k,j))^2));
      sum2   =   sum2+(sigma(j,k)^2);
  end
  ttheta(k)    =    acos(sqrt(sum1/sum2));
  totdeg(k)     =     ttheta(k)*360/(2*pi);
        end
%
%     Calculate  the  nyquist  array  elements
%
%      for  c  =  1:m
%   for  r  =  1:n
%        [retemp,imtemp]  =  nyquist(num(r
%
       clg
       subplot(221),loglog(w,sigma),xlabel('Frequency'),ylabel('Sing.Values')
       loglog(w,cond),xlabel('Frequency'),ylabel('Condition            Number')
       axis([x1   x2   0   90]);
       semilogx(w,deg),xlabel('Frequency'),ylabel('Alignment            (deg)')
       semilogx(w,totdeg),xlabel('Frequency'),ylabel('Interaction              (deg)')
       axis;
       end
FUNCTION [Aib,Bib,Cib,Wc2,Wo2,H,T] = intbal(A,B,C,dc)
%
% INTERNALLY-BALANCED
%                 State-space realization with internally balanced internal
%             coordinate system
%              ie.  Wc2 = Wo2 = S2       S2=second order modes
%
%               [Aib,Bib,Cib] = intbal(A,B,C,dc) returns a internally balanced
%                 state-space representation of the system (A,B,C)
%
%              [Aib,Bib,Cib,Wc2,Wo2,H,T] = intbal(A,B,C,dc) also returns matrices
%                 Wc2 and Wo2, the controllability and observability gramian,
%                   respectively, of the internally balanced realization,
%                matrix H, the Hankel matrix and matrix T, the simularity
%                transformation used to convert (A,B,C) to (Aon,Bon,Con)
%
%           Ref:
%                1) Moore, B.C., Principal Component Analysis in Linear
%                     Systems:  Controllability, Observability, and Model
%                  Reduction, IEEE Trans.AC, 26-1, Feb. 1981
%
%                2) Moore, B.C., Singular Value Analysis of Linear Systems,
%                  IEEE Decision and Control Conference, Jan. 1979, pp.66-73
%
%                3) Laub, A.J., "Computation of Balancing Transformations",
%                 Proc. JACC Conf., Vol. 1, paper FA8-E, 1980
%
%         see: balreal.m, inpnorm.m, outnorm.m
%
%
%    Variables:
%
%      A,B,C          = Original state space realization
%       Aib,Bib,Cib  = Balanced state space realization
%     dc              = Flag indicates the type of system:
%                          dc = 1, discrete time
%                          dc = 2, continuous time
%     Wc2             = Controllability gramian
%     Wo2            = Observability gramian
%     H             = Hankel matrix
%     T              = Transformation matrix
%
%
     j  =  sqrt(-1);
%
%    Calculate the observability and controllability gramians (continuous time)
%
%    a) Discrete case:
%
     if dc == 1
        c  =  ctrb(A,B);
        o  =  obsv(A,C);
        wc2  =  c*c';
```

```
            wo2 = o'*o;
%
%      b) continous case:
%
         else
            wc2 = gram(A,B);
            wo2 = gram(A',C');
         end
%
% Find the singular value decomposition of the gramians.  Note: the right and
% left singular vectors of the gramians are equal to the eigenvectors and the
% singular values of the gramians are equal to their respective eigenvalues
%
         [uc,sc,vc]  =  svd(wc2);
         ssc  =  sqrt(sc);
         [uo,so,vo]  =  svd(wo2);
         sso  =  sqrt(so);
%
% Find the Hankel matrix which describes the system and determine the
% second order modes (the singular values of H) Also, calculate the
% transformation matrix T.  Note: the second order modes are invariant under
%  internal coordinate transformations
%
         H  =  sso*vo'*vc*ssc;
         [uh,sh,vh]  =  svd(H);
         ssh  =  sqrt(sh);
         T  =  vo*(ssoh)*ssh;
%
%  The internally balanced state space realization
%
         Aib = T*T;
         Bib = T;
         Cib = C*T;
         if nargout > 3
            if dc == 1
               c  =  ctrb(Aib,Bib);
               o  =  obsv(Aib,Cib);
               Wc2 = c*c';
               Wo2 = o'*o;
            else
               Wc2 = gram(Aib,Bib);
               Wo2 = gram(Aib',Cib');
            end
         end
         mc  =  diag(diag(sh));
         mo  =  diag(diag(sh));
%        if mc ~= Wc2
%                 fprintf('Internal Balancing Incorrect')
%           mc
%           Wc2
%        end
%        if mo ~= Wo2
%                 fprintf('Internal Balancing Incorrect')
%           mo
%         Wo2
%      end
       end
```

```
FUNCTION [Ar,Br,Cr,k,hratio]=onmodred(A,B,C,minor,maxor,dc)
%
% OUTPUT-NORMAL MODEL REDUCTION:
%
%       Model Reduction scheme for a state space system description using
%       output normalization of the system (A,B,C).  Weak subsystems are
%       eliminated to produce a reduced order model (Ar,Br,Cr)
%
%       [Ar,Br,Cr,k] = onmodred(A,B,C,minor,maxor,dc) returns a output normalized
%                      kth order state space model of the system (A,B,C)
%
%        [Ar,Br,Cr,k,hratio] = onmodred(A,B,C,minor,maxor,dc) also returns the
%                      ratio of the (k+1 to n) and the first k second order
%                      modes of the system, hratio
%
%      Ref:
%
%           1) Moore, B.C., "Principal Component Analysis in Linear Systems:
%              Controllability, Observability and Model Reduction", IEEE Trans
%            AC, 26-1, Feb. 1981
%
%           2) Moore, B.C., "Singular Value Analysis of Linear Systems",
%             IEEE Decision and Control Conference, Jan. 1979, pp. 66-73
%
%
%     Variables:
%
%       A,B,C          = Original state space realization
%        Aib,Bib,Cib   = Output normalized state space realization
%        Ar,Br,Cr      = Reduced order output normalized state space realization
%       dc             = Flag indicates the type of system
%                             dc = 1, discrete time
%                             dc = 2, continuous time
%       minor          = Minimum order of the realizations
%       maxor          = Maximum order of the realizations
%       k              = Order of the realization
%        hratio        = Ratio of the k+1 to n second order modes to the first
%                          k second order modes, an error criteria
%       Wc2            = Controllability gramian
%       Wo2            = Observability gramian
%
%
       j  =  sqrt(-1);
%
%       Obtain the output-normalized model
%
       [Aon,Bon,Con,Wc2,Wo2,H,T]  =  outnorm(A,B,C,dc);
       [uc,sc,vc]=svd(Wc2);
%
%       Calculate the second order modes
%
       [uh,sh,vh]  =  svd(H);
       dsh  =  diag(sh);
```

```
%
%       Find a representative kth order model for (Aon,Bon,Con)
%
        n  =  length(dsh);
        for k = minor,maxor
           x1  =  sh(1:k);
           x2  =  sh(k+1:n);
           for i = 1,k
              X1(i)  =  x1(i)^2;
           end
           ii = n - (k+1)
           for i = 1,ii
              X2(i)  =  x2(i)^2;
           end
           hratio  =  sqrt(sum(X1))/sqrt(sum(X2));
           Ar = uc'*Aon*uc;
           Br = uc'*Bon;
           Cr = Con*uc;
           if dc == 1
              c  =  ctrb(Ar,Br);
              o  =  obsv(Ar,Cr);
              Wc2 = c*c';
              Wo2 = o'*o;
           else
              Wc2 = gram(Ar,Br);
              Wo2 = gram(Ar',Cr');
           end
           mo = diag(dsh)
           fprint('Controllability gramian for reduced order model')
           Wc2
           fprint('Diagonal matrix of the second order modes squared')
           mo
        end
```

```
FUNCTION [Aon,Bon,Con,Wc2,Wo2,H,T] = outnorm(A,B,C,dc)
%
% OUTPUT-NORMAL  State-space realization with output normalized internal
%                  coordinate system
%                   ie.  Wc2 = S4, Wo2 = I  S2=second order modes
%
%          [Aon,Bon,Con] = outnorm(A,B,C,dc) returns a output normalized
%                  state-space representation of the system (A,B,C)
%
%          [Aon,Bon,Con,Wc2,Wo2,H,T] = outnorm(A,B,C,dc) also returns matrices
%                  Wc2 and Wo2, the controllability and observability gramian,
%                   respectively, of the normalized realization, matrix H, the
%                  Hankel matrix and matrix T, the simularity transformation
%                 used to convert (A,B,C) to (Aon,Bon,Con)
%
%         Ref:
%               1) Moore, B.C., Principal Component Analysis in Linear
%                   Systems:  Controllability, Observability, and Model
%                 Reduction, IEEE Trans.AC, 26-1, Feb. 1981
%
%                2) Moore, B.C., Singular Value Analysis of Linear Systems,
%                  IEEE Decision and Control Conference, Jan. 1979, pp.66-73
%
%                3) Laub, A.J., "Computation of Balancing Transformations",
%                 Proc. JACC Conf., Vol. 1, paper FA8-E, 1980
%
%         see:  balreal.m,  inpnorm.m,  intbal.m
%
%   Variables:
%
%     A,B,C           = Original state space realization
%     Aon,Bon,Con  = Output normalized state space realization
%    dc               = Flag indicates the type of system:
%                           dc = 1, discrete time
%                           dc = 2, continous time
%    Wc2              = Controllability gramian
%    Wo2              = Observability gramian
%    H                = Hankel matrix
%    T                = Transformation matrix
%
%
     j  =  sqrt(-1);
%
%    Calculate  the  observability  and  controllability  gramians  (continuous  time)
%
%     a)  discrete  case:
%
     if dc == 1
        c  =  ctrb(A,B);
        o  =  obsv(A,C);
        wc2  =  c*c';
        wo2  =  o'*o;
%
```

```
%    b) continous case:
%
     else
        wc2 = gram(A,B);
        wo2 = gram(A',C');
     end
%
% Find the singular value decomposition of the gramians.  Note: the right and
% left singular vectors of the gramians are equal to the eigenvectors and the
% singular values of the gramians are equal to their respective eigenvalues
%
     [uc,sc,vc]  =  svd(wc2);
     ssc  =  sqrt(sc);
     [uo,so,vo]  =  svd(wo2);
     sso  =  sqrt(so);
%
% Find the Hankel matrix which describes the system and determine the
% second order modes (the singular values of H) Also, calculate the
% transformation matrix T.  Note: the second order modes are invariant under
%  internal coordinate transformations
%
     H  =  sso*vo'*vc*ssc;
     [uh,sh,vh]  =  svd(H);
     ssh  =  sqrt(sh);
     T  =  uo*(ssoh);
%
% The output normalized state space realization
%
     Aon = T*T;
     Bon = T;
     Con = C*T;
     if nargout > 3
        if dc == 1
           c  =  ctrb(Aon,Bon);
           o  =  obsv(Aon,Con);
           Wc2  =  c*c';
           Wo2  =  o'*o;
        else
           Wc2 = gram(Aon,Bon);
           Wo2 = gram(Aon',Con');
        end
     end
     mc  =  diag(diag(sh^2));
     mo = eye(Wo2);
%   if mc ~= Wc2
%             fprintf('Output Normalization Incorrect')
%        mc
%        Wc2
%    end
%    if mo ~= Wo2
%             fprintf('Output Normalization Incorrect')
%        mo
%        Wo2
%    end
```

end

```
% RDGA.M
%
%  This program calculates Witcher and McAvoy (1977) relative
%  dynamic gain array, Tung and Edgar (1977) relative dynamic
%   gain array and Bristol's relative gain array
%
%        num
%        G(s) =  ---
%         den
%
%
%      Variables:
%
%  G       = transfer function matrix
%  Go      = Steady state gain matrix
%  m      = number of inputs
%  n      = number of outputs
%  s       = Laplace transform variable
%  w      = frequency
%  num    = transfer function numerator polynomial
%  den    = transfer function denominator polynomial
%  M      = Witcher and McAvoy relative dynamic gain matrix
%  Mt     = Tung and Edgar relative dynamic gain matrix
%  RGA    = Bristol relative gain array
%  Lambda= M matrix over frequency range arranged row by row
%  alpha = Mt matrix over frequency range arranged row by row
%
      i  =  sqrt(-1);
%
%  Input the transfer function matrix
%
      m = input('Number of inputs, m = ')
      n = input('Number of outputs, n = ')
      fprintf('Input the elements of the transfer function matrix')
      fprintf('row by row starting with G(1,1) - G(1,m) and ending')
      fprintf('with  G(n,1)  -  G(n,m)')
      num = [];
      den  = [];
      for k = 1:n
  for j = 1:m
    temp = input('num = ')
    num = [num
    temp];
    temp = input('den = ')
    den = [den
    temp];
  end
      end
      clear  temp
%
%  Calculate the steady state gain matrix and the RGA
%
      j = 1;
```

```
        for r = 1:n
    for cc = 1:m
       Go(r,cc)      =      polyval(num(j,:),0)/polyval(den(j,:),0);
       j = j+1;
    end
        end
        xo = inv(Go);
        RGA = Go.*xo';
%
%   Input the frequency range of interest
%
        x1 = input('minimum frequency is 10 to the power x1 = ')
        x2 = input('maximum frequency is 10 to the power x2 = ')
        w = logspace(x1,x2);
        s = i*w;
        for k = 1:50
%
%       Calculate the polynomial TFM at s=i*w
%
   j = 1;
   for r = 1:n
      for cc = 1:m
         G(r,cc)      =      polyval(num(j,:),s(k))/polyval(den(j,:),s(k));
         j = j+1;
      end
   end
%
%  Calculate the RDGA's and place the elements of the RDGA's into
%   array's containing the information for the entire frequency range
%
   x = inv(G);
   j = 1;
   for r = 1:n
      for cc = 1:m
         M(r,cc)    =   G(r,cc)*x(cc,r);
         Mt(r,cc)    =   G(r,cc)*xo(cc,r);
         lambda(k,j)  =   M(r,cc);
         alpha(k,j)   =   Mt(r,cc);
         j = j+1;
      end
   end
       end
```

```
% SCALE.M
%
% SCALING - Geometric and Equilibration
%
%
%       This program scales a mxn matrix, G.  Geometric scaling involves
%       scaling the rows and the columns of the matrix by their respective
%       geometric means. Equilibration of the rows and the columns
%      involves scaling such that the maximum entry in each row and column
%      is of comparable magnitude
%
%   Variables:
%
%     G      = Unscaled matrix
%     Gs     = Scaled matrix
%    m      = Number of rows in G
%    n      = Number of columns in G
%      sfrow = Vector of row scaling factors
%      sfcol = Vector of column scaling factors
%      row    = Diagonal matrix of row scaling factors
%      col    = Diagonal matrix of column scaling factors
%
%       Geometric Scaling (Gill, Murray, Wright, "Process Optimization",
%   Academic Press, 1981, pp.353)
%
%  Compute the greatest ratio of two elements in the same column
%
%     G = g;
      diff=0.1;
      tol  =  0.0;
%      while diff >= tol
  max = 0.0;
  for j = 1:n
    for i = 1:m
      for k = 1:m
   if i ~= k
      if abs(G(k,j)) ~= 0
  ratio   =   abs(G(i,j)/G(k,j));
  if ratio > max
     max = ratio;
  end
      end
   end
       end
    end
  end
%
% Perform row scaling
%
  for i = 1:m
     minr  =  abs(G(i,1));
     maxr  =  abs(G(i,1));
     for j = 1:n
```

```
        if  abs(G(i,j))  <  minr
     if  abs(G(i,j))  ~=  0
        minr  =  abs(G(i,j));
     end
        elseif  abs(G(i,j))  >  maxr
     maxr  =  abs(G(i,j));
        end
      end
      sfrow(i)  =  (minr*maxr)^0.5;
      if  sfrow(i)  ~=  0
         Gr(i,1:n)    =    G(i,1:n)/sfrow(i);
      else
         Gr(i,1:n)    =   G(i,1:n);
      end
   end
   row  =  diag(sfrow);
%
%  Perform column scaling
%
   for  j  =  1:n
       minc   =   abs(Gr(1,j));
       maxc   =   abs(Gr(1,j));
       for  i  =  1:m
   if   abs(Gr(i,j))   <   minc
       if   abs(Gr(i,j))   ~=   0
          minc   =   abs(Gr(i,j));
       end
   elseif   abs(Gr(i,j))   >   maxc
       maxc   =   abs(Gr(i,j));
   end
         end
         sfcol(j)    =    (minc*maxc)^0.5;
         Gs(1:m,j)     =    Gr(1:m,j)/sfcol(j);
   end
   col    =    diag(sfcol);
%
%   Compute the greatest ratio of two elements in the same column
%
     max1  =  0.0;
     for  j  =  1:n
        for  i  =  1:m
   for  k  =  1:m
       if  i  ~=  k
          if   abs(Gs(i,j))   ~=   0
     ratio1    =    abs(Gs(i,j)/Gs(k,j));
     if  ratio1  >  max1
        max1  =  ratio1;
     end
        end
      end
   end
       end
     end
%
```

```
%  Check for termination
%
  diff  =  abs(max1-max)
  tol  =  0.1*abs(max)
%  if  diff  >=  tol
%     G  =  Gs;
%      flag  =  1
% end
%     end
%
%
%  Perform  column  scaling  followed  by  row  scaling
%
%
% Perform  column  scaling
%
%     G  =  g;
      diffa=0.1;
      tola  =  0.0;
%       while  diff  >=  tol
  for  j  =  1:n
      minc  =  abs(G(1,j));
      maxc  =  abs(G(1,j));
      for  i  =  1:m
  if  abs(G(i,j))  <  minc
      if  abs(G(i,j))  ~=  0
         minc  =  abs(G(i,j));
      end
  elseif  abs(G(i,j))  >  maxc
      maxc  =  abs(G(i,j));
  end
       end
       sfcol(j)  =  (minc*maxc)^0.5;
       Gc(1:m,j)  =  G(1:m,j)/sfcol(j);
  end
  cola  =  diag(sfcol);
%
% Perform  row  scaling
%
  for  i  =  1:m
     minr  =  abs(Gc(i,1));
     maxr  =  abs(Gc(i,1));
     for  j  =  1:n
         if  abs(Gc(i,j))  <  minr
    if  abs(Gc(i,j))  ~=  0
       minr  =  abs(Gc(i,j));
    end
         elseif  abs(Gc(i,j))  >  maxr
    maxr  =  abs(Gc(i,j));
         end
     end
     sfrow(i)  =  (minr*maxr)^0.5;
     if  sfrow(i)  ~=  0
         Gsa(i,1:n)  =  Gc(i,1:n)/sfrow(i);
     else
        Gsa(i,1:n)  =  Gc(i,1:n);
     end
  end
  rowa = diag(sfrow);
%
%  Compute the greatest ratio of two elements in the same column
%
   maxa = 0.0;
   for j = 1:n
      for i = 1:m
  for k = 1:m
     if i ~= k
        if  abs(Gsa(i,j))  ~=  0
    ratiola   =   abs(Gsa(i,j)/Gsa(k,j));
    if ratiola > maxa
       maxa = ratiola;
    end
        end
     end
  end
      end
   end
%
%  Check for termination
%
  diffa= abs(maxa-max)
  tola=  0.1*abs(max)
%  if diffa >= tola
%     G = Gsa;
%      flag = 1
% end
%     end
      end
```